CONSTRAINED ADAPTIVE LEAST MEAN SQUARED FILTERS

Michael H. Davis

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

CONSTRAINED ADAPTIVE LEAST MEAN SQUARED FILTERS

by

Michael H. Davis

June 1981

Thesis Advisor: S. R. Parker

Approved for public release; distribution unlimited.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Constrained Adaptive Least Mean Squared Filters		5. TYPE OF REPORT & PERIOD COVERED Master's thesis; June 1981
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Michael H. Davis		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		12. REPORT DATE June 1981
		13. NUMBER OF PAGES 125
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Adaptive Filters; Constrained Adaptive Filters; LMS

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The possibility of designing constrained adaptive finite impulse response digital filters is investigated as motivated by a study of adaptive noise cancellation. The first constraint considered consists of a fixed angle between filter zeros and is implemented in a master-slave approach in which one of the zeros is adjusted adaptively and the others follow subject to the constraint. The second constraint considered is a linear

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601 |

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(20. ABSTRACT Continued)

constraint on the filter weights and is implemented by augmenting the error equation with Lagrangean multipliers. Simulations indicate that the approach is feasible.

Constrained Adaptive Least Mean Squared Filters

by

Michael H. Davis
Lieutenant, United States Navy
B.S.E.E., University of Florida, 1972

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN ELECTRICAL ENGINEERING

from the

NAVAL POSTGRADUATE SCHOOL
June 1981

ABSTRACT

The possibility of designing constrained adaptive finite impulse response digital filters is investigated as motivated by a study of adaptive noise cancellation. The first constraint considered consists of a fixed angle between filter zeros and is implemented in a master-slave approach in which one of the zeros is adjusted adaptively and the others follow subject to the constraint. The second constraint considered is a linear constraint on the filter weights and is implemented by augmenting the error equation with Lagrangean multipliers. Simulations indicate that the approach is feasible.

TABLE OF CONTENTS

I. INTRODUCTION

This thesis investigates the possibility of designing constrained adaptive finite impulse response (FIR) filters. FIR adaptive filters self-adjust their parameters to match the output signal to a desired signal in an optimal least squares sense. Considered in detail is the possibility of designing the adaptive process so that either the parameters (weights) of the filter are constrained by an algebraic formula or, alternatively, are constrained so that the zero pattern of the FIR transfer function is fixed while the actual positions of the zeros are adaptively adjusted. An application which motivated this study is adaptive noise cancellation where noise rejection over a given bandwidth with a specified frequency rejection spectrum is required, but the center frequency of the rejection band is unknown and needs to be determined adaptively.

The usual method of extracting a signal from a strong noise background is to pass it through a filter which tends to suppress the noise while leaving the signal relatively unchanged. Filters designed for this purpose can be either fixed or adaptive. The design of fixed filters is based on some prior knowledge of both the signal and noise characteristics. Adaptive filters have the distinct advantage of being able to adjust themselves automatically and their implementation requires little a priori knowledge of the

signal and noise characteristics. This type of adaptive filtering often converges to the optimal filter, which originated with the pioneering work of Wiener [Ref. 1] and later was enhanced by Kalman [Ref. 2] and others. The optimal Wiener filter is defined as the linear filter optimized with respect to a minimum mean squared error, where the error is the difference between the filter output and the desired filter output.

Widrow [Ref. 3] presents the classic FIR adaptive filter which is optimized via a gradient minimum seeking algorithm called the Least Mean Squared (LMS) algorithm. This chapter discusses the theory behind both the Wiener and LMS filters, and demonstrates that, for statistically stationary input signals, the steady-state values of the LMS adaptive filter weights are accurate approximations of the Wiener weights.

The following matrix, vector, and scalar definitions are used in this thesis. An underlined capital letter denotes a matrix ($\underline{M}$). A lower case underlined letter denotes a vector ($\underline{v}$). A lower case letter which is not underlined denotes a scalar (s). Finally, a capital letter which is not underlined denotes an internal element of the corresponding matrix M(n).

In Chapter II the concept of adaptive noise cancelling is studied in detail. In Chapter III the constrained adaptive FIR filter is presented and in Chapter IV some simulation results are given.

A literature search has indicated that very little research has been done in the area of constrained adaptive filters. Frost [Ref. 8] presents a constrained LMS algorithm which is capable of adjusting an array of sensors in real time to respond to a signal coming from a desired direction while discriminating against noises coming from other directions. A set of linear constraints on the weights maintains a chosen frequency characteristic for the array in the direction of interest.

A. THE ADAPTIVE LMS AND THE OPTIMAL WIENER FILTERS

The LMS adaptive filter shown in Figure 1.1 uses the weighted sum of a set of input signals which are combined to form an output signal y(n). The input signal vector is defined as

$$\begin{aligned} \underline{X}(n) &= [X_1(n) \quad X_2(n) \quad \ldots \quad X_N(n)]^T \\ &= [X(n) \quad X(n-1) \quad \ldots \quad X(n-N)]^T \end{aligned} \tag{1.1}$$

The weighting coefficients

$$\underline{W} = [W_1 \quad W_2 \quad \ldots \quad W_N]^T \tag{1.2}$$

are the weights of the system. Each input value is multiplied by a corresponding weight coefficient and the linear combination of the sum of these weighted inputs forms the output

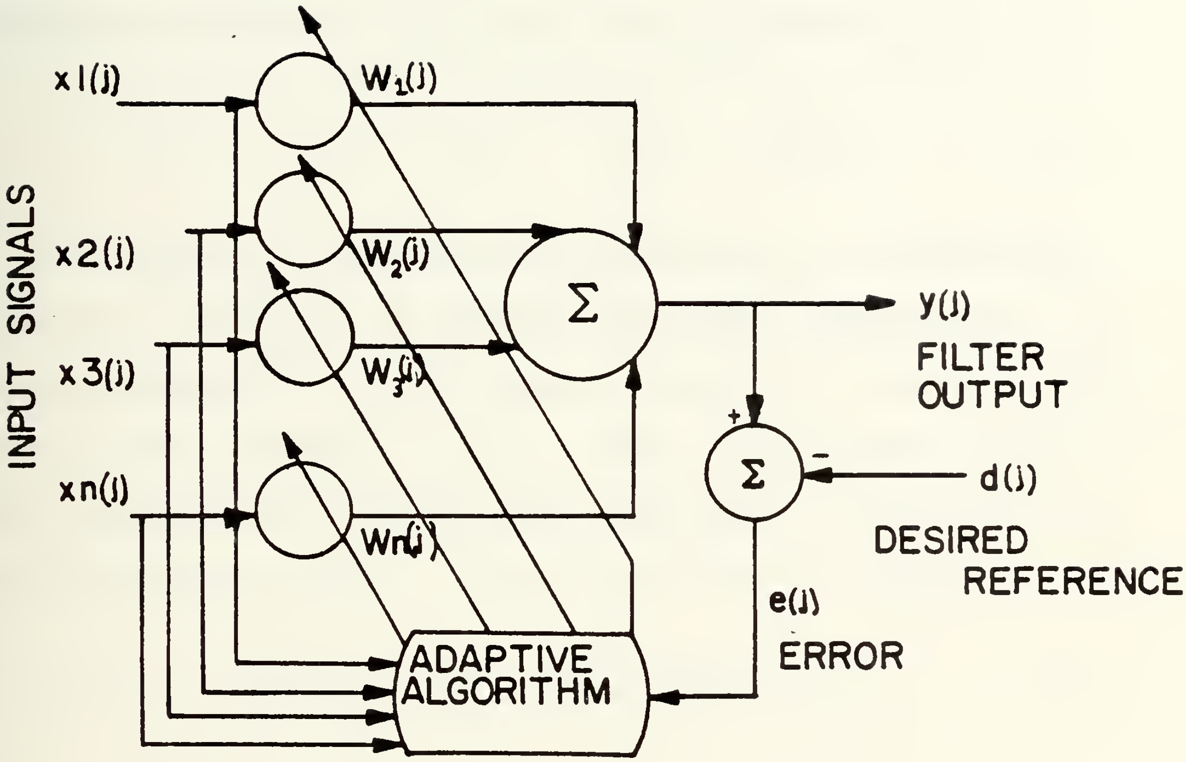

Fig. 1.1. The LMS Adaptive Filter

$$y(n) = \sum_{i=1}^{N} W_i X_i(n) = \sum_{i=1}^{N} W_i X(n-i) \qquad (1.3)$$

which can be written in matrix form as

$$y(n) = \underline{X}(n)^T \underline{W} = \underline{W}^T \underline{X}(n) \qquad (1.4)$$

An error signal is now defined as the difference between a desired response d(n) and the actual response y(n).

$$\varepsilon(n) = d(n) - y(n) = d(n) - \underline{W}^T \underline{X}(n) \qquad (1.5)$$

The purpose of the adaptive algorithm is to adjust the weights of the filter to minimize the mean-square error. A general expression for mean-square error as a function of the weight values, assuming that the input signals and the desired response are statistically stationary, can be derived in the following manner. Squaring the error results in

$$\varepsilon(n)^2 = d(n)^2 - 2d(n)\underline{X}(n)^T\underline{W} + \underline{W}^T\underline{X}(n)\underline{X}(n)^T\underline{W} \qquad (1.6)$$

Taking the expected value of both sides yields

$$E[\varepsilon(n)^2] = E[d(n)^2] - 2E[d(n)\underline{X}(n)^T]\underline{W} + \underline{W}^T E[\underline{X}(n)\underline{X}(n)^T] \qquad (1.7)$$

If the vector $\underline{r}_{xd}$ is now defined as the cross correlation between the desired response (a scalar value) and the input vector, the result is

$$\underline{r}_{xd} = E[d(n)\underline{X}(n)] = E[d(n)X(n) \quad d(n)X(n-1) \quad \ldots \quad d(n)X(n-N)] \tag{1.8}$$

The input correlation matrix $\underline{R}_{xx}$ is defined as

$$\underline{R}_{xx} = E[\underline{X}(n)\underline{X}(n)^T] = E\begin{bmatrix} X1(n)X1(n) & X1(n)X2(n) & \ldots & X1(n)XN(n) \\ X2(n)X1(n) & X2(n)X2(n) & \ldots & X2(n)XN(n) \\ \vdots & & & \\ XN(n)X1(n) & XN(n)X2(n) & \ldots & XN(n)XN(n) \end{bmatrix} \tag{1.9}$$

$$= E\begin{bmatrix} X(n)^2 & X(n)X(n-1) & \ldots & X(n)X(n-N) \\ X(n-1)X(n) & X(n-1)^2 & \ldots & X(n-1)X(n-N) \\ \vdots & & & \\ X(n-N)X(n) & X(n-N)X(N-1) & \ldots & X(n-N)^2 \end{bmatrix} \tag{1.10}$$

Equation (1.10) is usually written as:

$$\underline{R}_{xx} = \begin{bmatrix} R_{xx}(0) & R_{xx}(1) & \ldots & R_{xx}(N) \\ R_{xx}(1) & R_{xx}(0) & \ldots & R_{xx}(N-1) \\ \vdots & & & \\ R_{xx}(N) & R_{xx}(N-1) & \ldots & R_{xx}(0) \end{bmatrix} \tag{1.11}$$

Now the mean squared error (1.7) can be expressed as

$$E[\varepsilon(n)^2] \quad = \quad E[d(n)^2] - 2\underline{r}_{xd}\underline{W} + \underline{W}^T\underline{R}_{xx}\underline{W} \qquad (1.12)$$

Since the error is a quadratic function of the weights and $\underline{R}_{xx}$ is a positive definite matrix then the surface of the error has a guaranteed global minimum. Gradient methods adjust the weights to minimize the error by descending along this surface with the objective of finding the bottom.

The gradient $\underline{\nabla}$ of the quadratic error function is obtained by differentiating Equation (1.12) with respect to the weight vector $\underline{w}$.

$$\underline{\nabla} \quad = \quad \left\{\frac{\partial E[\varepsilon(n)^2]}{\partial w_1} \quad \frac{\partial E[\varepsilon(n)^2]}{\partial w_2} \quad \cdots \quad \frac{\partial E[\varepsilon(n)^2]}{\partial w_N}\right\}^T \qquad (1.13)$$

$$\underline{\nabla} \quad = \quad -2\underline{r}_{xd} + 2\underline{R}_{xx}\underline{w} \qquad (1.14)$$

This optimal weight vector $\underline{w}$ is called the Wiener weight vector [Ref. 1] and is found by setting the gradient of the mean square error function to zero.

$$\underline{W}^* \quad = \quad \underline{R}_{xx}^{-1}\ \underline{r}_{xd} \qquad (1.15)$$

This equation is a matrix version of the Wiener-Hopf equation. The practical objective of the adaptive system is to find a solution to (1.15). An exact solution would require a priori knowledge of the correlation matrices $\underline{r}_{xd}$ and $\underline{R}_{xx}$. However, this information is usually not available. Additionally, when the number of weights is large a direct

solution is computationally cumbersome since it requires an N by N matrix inversion in addition to the N(n+1)/2 auto-correlation and cross correlation measurements.

B. AN EXAMPLE OF THE WIENER SOLUTION

As an example to illustrate the calculations involved for a simple four weight Wiener solution consider the following example using deterministic signals. The input signal is the sampled sum of two sinusoids of different frequencies and is given by

$$X(nT) = \sin(\omega_1 nT) + \sin(\omega_2 nT) \qquad (1.16)$$

The desired value is the sampled desired or reference signal.

$$d(nT) = K\sin(\omega_2 nT) \qquad (1.17)$$

In order to calculate the autocorrelation matrix, $\underline{\underline{R}}_{xx}$, note that

$$\begin{aligned}
X1(n) &= X(n) = \sin(\omega_1 nT) + \sin(\omega_2 nT) \\
X2(n) &= X(n-1) = \sin(\omega_1 (n-1)T) + \sin(\omega_2 (n-1)T) \\
X3(n) &= X(n-2) = \sin(\omega_1 (n-2)T) + \sin(\omega_2 (n-2)T) \\
X4(n) &= X(n-3) = \sin(\omega_1 (n-3)T) + \sin(\omega_2 (n-3)T)
\end{aligned} \qquad (1.18)$$

The following expected value computations are taken over a full number of cycles, P, for both sinusoids.

$$E[X1(n)X1(n)] = E[X1(n)^2] = \overline{X1(n)^2}$$

$$= \frac{1}{2} + \frac{1}{2} = 1$$

Likewise all the diagonal terms will take on the value of

$$\overline{X2(n)^2} = \overline{X3(n)^2} = \overline{X4(n)^2} = 1$$

Consider now the off diagonal terms of $\underline{R}_{\underline{xx}}$

$$E[X1(n)X2(n)] = \overline{X1(n)X2(n)}$$

$$= \sum_{n=0}^{P} (\sin \omega_1 nT + \sin \omega_2 nT)(\sin\omega_1 (n-1)T + \sin\omega_2 (n-1)T)$$

Carrying out the indicated multiplication yields

$$E[X1(n)X2(n)] = \sum_{n=0}^{P} (\sin\omega_1 nT)(\sin\omega_1 (n-1)T)$$

$$+ \sum_{n=0}^{P} (\sin\omega_2 nT)(\sin\omega_2 (n-1)T)$$

$$+ \sum_{n=0}^{P} (\sin\omega_1 nT)(\sin\omega_2 (n-1)T)$$

$$+ \sum_{n=0}^{P} (\sin\omega_2 nT)(\sin\omega_1 (n-1)T)$$

The final two terms of this expression are zero leaving

$$E[X1(n)X2(n)] = \sum_{n=0}^{P} (\sin\omega_1 nT \sin(\omega_1 nT-\omega_1 T)+\sin\omega_2 nT \sin(\omega_2 nT-\omega_2 T))$$

Using the identity $(\sin A)(\sin B) = \frac{1}{2}[\cos(A-B)-\cos(A+B)]$ the final expression becomes

$$E[X1(n)X2(n)] = \frac{1}{2}\cos\omega_1 T + \frac{1}{2}\cos\omega_2 T \tag{1.19}$$

Similarly,

$$E[X1(n)X3(n)] = \frac{1}{2}\cos 2\omega_1 T + \frac{1}{2}\cos 2\omega_2 T \tag{1.20}$$

$$E[X1(n)X4(n)] = \frac{1}{2}\cos 3\omega_1 T + \frac{1}{2}\cos 3\omega_2 T \tag{1.21}$$

Denoting

$$\begin{aligned} E[X1(n)X2(n)] &= a \\ E[X1(n)X3(n)] &= b \\ E[X1(n)X4(n)] &= c \end{aligned} \tag{1.22}$$

the autocorrelation matrix can be expressed as

$$\underline{\underline{R_{xx}}} = \begin{bmatrix} 1 & a & b & c \\ a & 1 & a & b \\ b & a & 1 & a \\ c & b & a & 1 \end{bmatrix} \tag{1.23}$$

For the cross correlation matrix $\underline{r}_{xd}$

$$\underline{r}_{xd} = E[d(n)X1(n) \quad d(n)X2(n) \quad d(n)X3(n) \quad d(n)X4(n)]^T \tag{1.24}$$

These terms compute as

$$E[d(n)X1(n)] = \overline{K(\sin\omega_1 nT + \sin\omega_2 nT)\sin\omega_1 nT}$$

$$= K \sum_{n=0}^{P} \sin\omega_1 nT(\sin\omega_1 nT + \sin\omega_2 nT)$$

$$= K \sum_{n=0}^{P} \sin^2\omega_1 nT + K \sum_{n=0}^{P} (\sin\omega_1 nT)\sin\omega_2 nT$$

The final term sums to zero leaving

$$E[d(n)X1(n)] = K \sum_{n=0}^{P} \sin^2\omega_1 nT = \frac{K}{2} \tag{1.25}$$

Similarly,

$$E[d(n)X2(n)] = \overline{K(\sin\omega_1 (n-1)T + \sin\omega_2 (n-1)T)\sin\omega_1 nT}$$

$$= K \sum_{n=0}^{P} \sin\omega_1 nT(\sin\omega_1 (n-1)T)$$

$$+ K \sum_{n=0}^{P} \sin\omega_1 nT(\sin\omega_2 (n-1)T)$$

The final term sums to zero leaving

$$E[d(n)X2(n)] = K[\sum_{n=0}^{P} \sin^2\omega_1 nT \cdot \sum_{n=0}^{P} \sin\omega_1 nT\sin\omega_1 T]$$

Using the (sinA)(sinB) trigonometric identity invoked for (1.19) and realizing that the final term sums to zero, the expected value finally becomes

$$E[d(n)X2(n)] = \frac{K}{2}\cos\omega_1 T \tag{1.26}$$

Likewise,

$$E[d(n)X3(n)] = \frac{K}{2}\cos 2\omega_1 T \tag{1.27}$$

and

$$E[d(n)X4(n)] = \frac{K}{2}\cos 3\omega_1 T \tag{1.28}$$

The Wiener solution for $\underline{w}^*$ as given by (1.15) is

$$\underline{W}^* = \begin{bmatrix} 1 & a & b & c \\ a & 1 & a & b \\ b & a & 1 & a \\ c & b & a & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ \cos\omega_1 T \\ \cos 2\omega_1 T \\ \cos 3\omega_1 T \end{bmatrix} \frac{K}{2} \tag{1.29}$$

where the values of a, b and c are given by (1.22). The Wiener vector is calculated for the specific values of

$K = 1$

$f1 = 10$ Hz

$f2 = 35$ Hz

$T = 1/256$

For these specific values

$a = .8116$

$b = .3676$

$c = -.052$

$\frac{K}{2}[\cos\omega_1 T] = .485$

$\frac{K}{2}[\cos\omega_1 T] = .441$

$\frac{K}{2}[\cos\omega_1 T] = .3705$

and the optimal Wiener vector $\underline{w}^*$ is

$$\underline{w}^* = \begin{bmatrix} w_1^* \\ w_2^* \\ w_3^* \\ w_4^* \end{bmatrix} = \begin{bmatrix} 3.0614 \\ -5.5772 \\ 5.1228 \\ -1.5780 \end{bmatrix} \tag{1.31}$$

C. THE LMS ADAPTIVE FIR FILTER

The values for the correlation matrices are not generally known a priori. The LMS adaptive algorithm introduced by Widrow and Hoff [Ref. 1] is a practical method for finding close approximate solutions to (1.15) in real time. The algorithm does not require measurements of correlation functions, nor does it require matrix inversion. The LMS algorithm is an implementation of the method of steepest

descent. According to this method, the "next" weight w(n+1) is equal to the present weight w(n) plus a change proportional to the negative gradient. Thus

$$w(n+1) = w(n) - \mu\nabla(n)$$

The parameter μ controls stability and also the rate of convergence. An estimate of the instantaneous gradient $\nabla(n)$ is obtained by assuming that the square of a single error sample $\varepsilon(n)$ is an estimate of the mean square error. Differentiating $\varepsilon(n)$ with respect to w results in

$$\hat{\nabla}(n) = \left[\frac{\partial \varepsilon(n)^2}{\partial w_1} \quad \frac{\partial \varepsilon(n)^2}{\partial w_2} \quad \cdots \quad \frac{\partial \varepsilon(n)^2}{\partial w_n}\right]^T \tag{1.32}$$

$$\hat{\nabla}(n) = 2\,\varepsilon(n)\left[\frac{\partial \varepsilon(n)}{\partial w_1} \quad \frac{\partial \varepsilon(n)}{\partial w_2} \quad \cdots \quad \frac{\partial \varepsilon(n)}{\partial w_N}\right]^T \tag{1.33}$$

The expression for the gradient estimate can be approximated by

$$\hat{\nabla}(n) = -2\varepsilon(n)X(n) \tag{1.34}$$

Using this estimate in place of the true gradient yields the Widrow-Hoff LMS algorithm given by

$$w(n+1) = w(n) + 2\mu\varepsilon(n)X(n) \tag{1.35}$$

The algorithm is generally easy to implement and although it makes use of gradients of the mean square error

function, it does not require squaring, averaging or differentiation.

To show the convergence of the Widrow-Hoff LMS algorithm (1.35) to the Wiener solution given by (1.15) write

$$\underline{w}(n+1) = \underline{w}(n) + 2\mu\varepsilon(n)\underline{X}(n)$$

as

$$\underline{w}(n+1) = \underline{w}(n) + 2\mu\underline{X}(n)[d(n) - \underline{X}(n)^T\underline{w}(n)]$$

$$\underline{w}(n+1) = [I - 2\mu\underline{X}(n)\underline{X}(n)^T]\underline{w}(n) + 2\mu\underline{X}(n)d(n)$$

Consider now the ensemble average. That is,

$$E[\underline{w}(n+1)] = [I - 2\mu\underline{\underline{R}}_{xx}]E[\underline{w}(n)] + 2\mu\underline{r}_{xd} \tag{1.36}$$

With an initial weight vector $\underline{w}(0)$, j+1 iterations of Equation (1.36) becomes

$$E[\underline{w}(j+1)] = [I-2\mu\underline{\underline{R}}_{xx}]^{j+1}\underline{w}(0) + 2\mu\sum_{i=0}^{j}[I-2\mu\underline{\underline{R}}_{xx}]^{i}\underline{r}_{xd} \tag{1.37}$$

This equation may be put in diagonal form by using the normal form expansion of the matrix $\underline{\underline{R}}_{xx}$; that is

$$\underline{\underline{R}}_{xx} = Q^{-1}\Lambda Q$$

where Λ is the diagonal matrix of eigenvalues, and the square matrix of eigenvectors is the matrix Q. Equation (1.37) can now be written as

$$[[w(j+1)] = [I-2\mu Q^{-1}\Lambda Q]^{j+1}\underline{w}(0) + 2\mu \sum_{i=0}^{j} [I-2\mu Q^{-1}\Lambda Q]^{i}\underline{r}_{\underline{x}d}$$

$$= Q^{-1}[I-2\mu\Lambda]^{j+1}Q\underline{w}(0) + 2\mu \sum_{i=0}^{j} [I-2\mu Q^{-1}\Lambda Q]^{i}\underline{r}_{\underline{x}d} \qquad (1.39)$$

As long as the terms of the diagonal matrix $[I-2\mu\Lambda]$ are all of magnitude less than unity, then the first term of (1.39) goes to zero as the number of iterations increases. That is,

$$\lim_{j\to\infty}[I-2\mu\Lambda]^{j+1} \to 0 \qquad (1.40)$$

Writing (1.40) as a geometric series results in

$$\lim_{j\to\infty} \sum_{i=0}^{j} [I-2\mu\Lambda]^{i} = \frac{1}{2\mu}\Lambda^{-1} \qquad (1.41)$$

or for the specific component 'p' of the matrix

$$\lim_{j\to\infty} \sum_{i=0}^{j} [I-2\mu\Lambda]^{i} = \frac{1}{2\mu\lambda_p}$$

Therefore, in the limit, Equation (1.39) becomes

$$\lim_{j\to\infty} E[\underline{w}(j+1)] = Q^{-1}\Lambda Q\underline{r}_{\underline{x}d} = \underline{R}_{\underline{xx}}^{-1}\,\underline{r}_{\underline{x}d} \qquad (1.42)$$

which is the same as the Wiener solution shown in Equation (1.15).

Convergence of the mean of the weight vector to the Wiener solution is insured if and only if the proportionality constant μ is set within certain bounds. Since the diagonal terms of $[I-2\mu\Lambda]$ must all have magnitude less than unity, and since all eigenvalues in Λ are positive, the bounds on μ are given by

$$|1 - 2\mu\lambda_{max}| < 1$$

or

$$0 < \mu < 1/\lambda_{max} \tag{1.43}$$

where λ_{max} is the maximum eigenvalue of $\underline{R}_{\underline{\underline{xx}}}$.

II. ADAPTIVE NOISE CANCELLING AND ITS APPLICATIONS

In this chapter the concept of adaptive noise cancellation is considered. Adaptive noise cancelling is one of the most practical applications of adaptive signal processing [Refs. 4,5,6,7]. The basic principle involved is the use of a reference input derived from one or more sensors located at points in the noise field where the signal is either undetectable or weak. The reference input is adaptively filtered and subtracted from the primary input containing both signal and noise to generate an error signal which controls the adaptive process. The result is the attenuation or elimination of the primary noise by cancellation. In circumstances where adaptive noise cancelling is applicable levels of noise rejection are often attainable that would be difficult or impossible to achieve through direct filtering. Because the concepts of adaptive noise cancelling and their applications are fundamental to constrained adaptive filtering developed in Chapter III, they are presented here in detail. For example, if the angle between zeros and the magnitude of the zeros are maintained constant during an adaptive process, the frequency response characteristics remains invariant. Specific examples are discussed and the resulting equations indicate the design limitations of these approaches.

A. THE BASIC SYSTEM

Figure 2.1 depicts the basic adaptive noise cancelling system concept . A signal [s] is transmitted over a channel

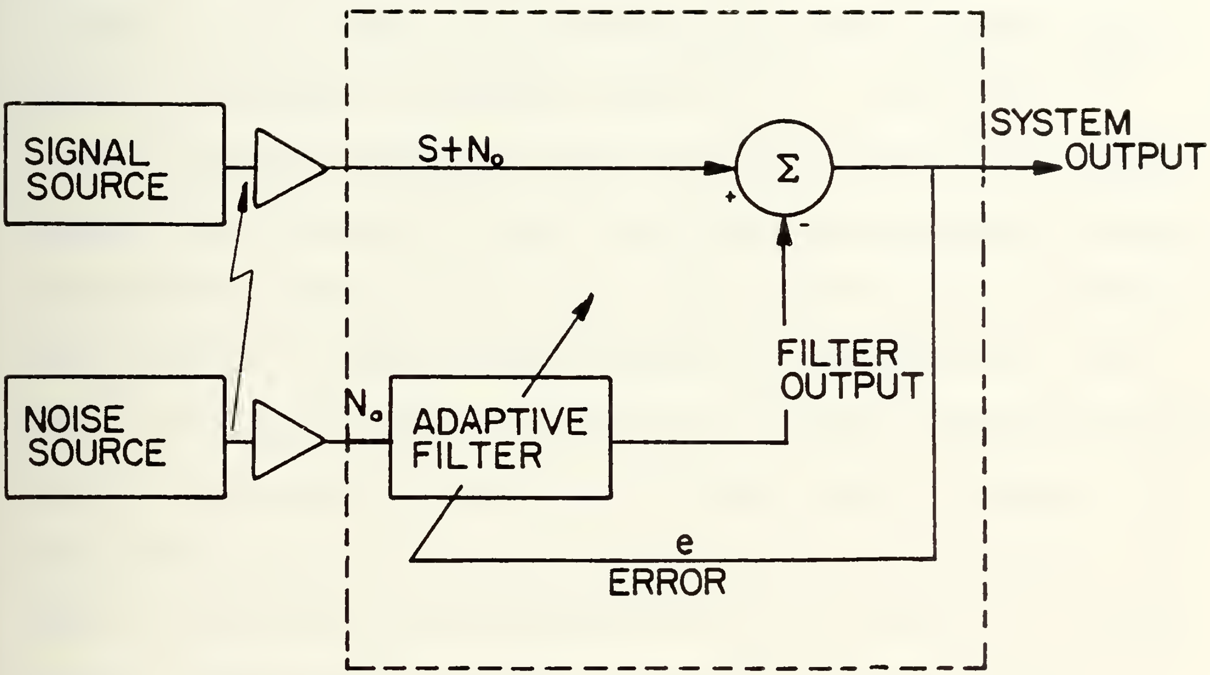

Fig. 2.1. The Adaptive Noise Cancelling Concept

to a sensor that picks up a noise [n0] which us uncorrelated with the signal. The combined signal and noise [s+n0] form the primary input to the canceller. A second sensor receives a noise [n1] which is also uncorrelated with the signal but which is correlated in some unknown manner with the noise [n0]. This sensor input provides the reference input to the canceller. The noise [n1] is filtered to produce an output [y] that is as close a replica as possible to [n0]. This output is subtracted from the primary input [s+n0] to produce the system output [z = s + n0 - y].

If it were possible to know the characteristics of the channels over which the noise was transmitted to the primary and reference sensors, then it would theoretically be possible to design a fixed filter capable of changing [n1] into [n0]. The filter output could then be merely subtracted from the primary input to produce the signal alone. However, the characteristics of the transmission paths are generally unknown or known only approximately and are seldom of a fixed nature, therefore precluding the use of a fixed filter.

In the system shown in Figure 2.1 the reference input is processed by an adaptive filter. Self-adjustment of the filter's impulse response is accomplished by the LMS algorithm described in Chapter I. The error signal used in the adaptive process depends on the nature of the application. With noise cancelling systems the practical objective is to produce a system output [z = s + n0 - y] that is a best fit in the least squares sense to the signal [s].

Consider the system inputs shown in Figure 2.1 and the filter output [y]. Assume that [s] is uncorrelated with [n0] and [n1] and that [n0] and [n1] are correlated. The output [z] is

$$z = s + n0 - y \tag{2.1}$$

Squaring produces

$$z^2 = s^2 + (n0-y)^2 + 2s(n0-y) \tag{2.2}$$

Now, taking the expectation of both sides and using the fact that [s] is uncorrelated with [n0] and [y] produces

$$\begin{aligned} E[z^2] &= E[s^2] + E[(n0-y)^2] + 2E[s(n0-y)] \\ &= E[s^2] + E[(n0-y)^2] \qquad (2.3) \end{aligned}$$

The minimum output power is

$$\text{Min } E[z^2] = E[s^2] + \text{Min } E[(n0-y)^2] \tag{2.4}$$

The signal power is unaffected as the filter is adjusted to minimize $E[z^2]$. When the filter is adjusted so that $E[z^2]$ is minimized, $E[(n0-y)^2]$ is therefore also minimized. The filter output [y] is then a best least squares estimate of the primary noise [n0]. It is also of interest to note that when $E[(n0-y)^2]$ is minimized, $E[(z-s)^2]$ is also minimized, since from (2.1)

$$(z-s) = (n0-y) \qquad (2.5)$$

Adapting the filter to minimize the total output power causes the output [z] to be a best least squares estimate of the signal [s] for the given reference input. The output [z] will contain the signal [s] plus noise. From (2.1) the output noise is given by (n0-y). Since minimizing $E[z^2]$ minimizes $E[(n0-y)^2]$, minimizing the total output power minimizes the output noise power. Because the output signal remains constant, minimizing the total output power maximizes the signal-to-noise ratio. From (2.3) it can be seen that the smallest possible output power $E[z^2] = E[s^2]$ is achieved when $E[(n0-y)^2] = 0$ and therefore $y = n0$ and $z = s$. In this case, minimizing output power causes the output signal to be perfectly noise free. These same arguments can be extended to the case where the primary and reference inputs contain, in addition to [n0] and [n1], additive random noises uncorrelated with each other and with [s], [n0] and [n1]. They can also be extended to the case where [n0] and [n1] are deterministic rather than stochastic.

B. SIGNAL-TO-NOISE RATIO IN THE ADAPTIVE FILTER

At this point it is of value to show analytically the increase in signal-to-noise ratio of the noise cancelling technique.

As noted previously, fixed filters are generally inappropriate for noise cancelling situations because the correlation

and cross correlation functions are usually unknown and additionally they often vary with time. Adaptive filters "learn" these statistics initially and then track them through slow variations. For stationary stochastic inputs, however, the steady-state performance of adaptive filters closely approximates that of fixed Wiener filters and therefore Wiener filter theory provides a useful mathematical tool in carefully analyzing statistical noise cancelling problems.

Figure 2.2 shows the classic single-input/single-output Wiener filter. The input signal is $x(j)$, the output signal is $y(j)$ and the desired response is $d(j)$. The input and output signals are assumed to be discrete in time, and the input signal and desired response are assumed to be statistically stationary. The error signal is $e(j) = d(j) - y(j)$. The filter is linear, discrete and designed to be optimal in the minimum mean-square-error sense. It is considered to be composed of an infinitely long, two sided tapped delay line.

As shown in Chapter I (1.15) the optimal Wiener weight solution can be written

$$\underline{w}^*(k) = \underline{\underline{R}}_{xx}^{-1}(k)\,\underline{r}_{xd}(k) \tag{2.6}$$

or

$$\sum_{\ell=-\infty}^{\infty} \underline{w}^*(\ell)\,\underline{\underline{R}}_{xx}(k-\ell) = \underline{r}_{xd}(k) \tag{2.7}$$

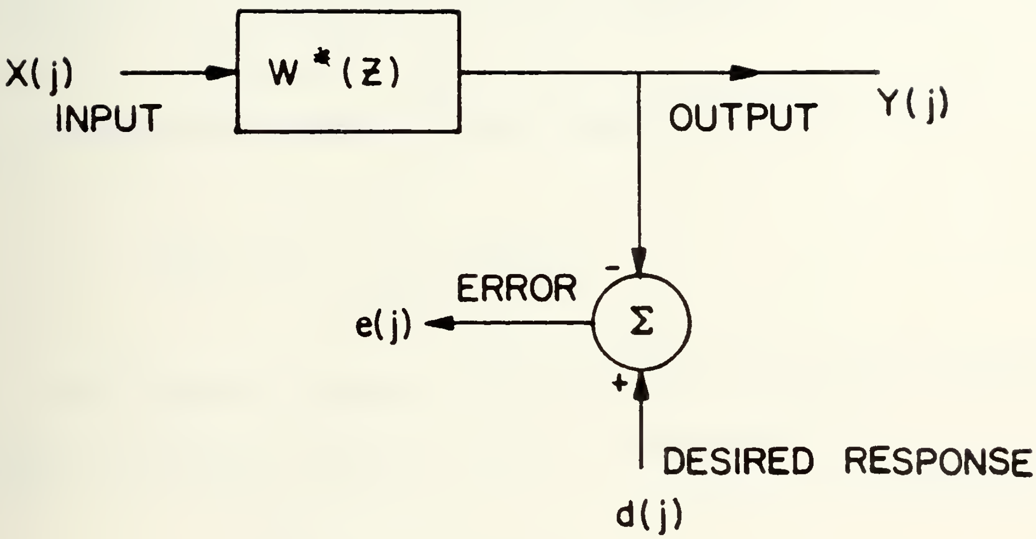

Fig. 2.2. The Wiener Filter

To obtain the transfer function of the Wiener filter consider first the power spectral density of the process. The application of the z-transform to $\underline{R}_{xx}(k)$ yields

$$S_{xx}(z) \triangleq z[\underline{R}_{xx}(k)] = \sum_{k=-\infty}^{\infty} \underline{R}_{xx}(k) z^{-k} \tag{2.8}$$

Likewise, the cross power spectrum between the input signal and desired response is

$$S_{xd}(z) \triangleq z[\underline{r}_{xd}(k)] = \sum_{k=-\infty}^{\infty} \underline{r}_{xd}(k) z^{-k} \tag{2.9}$$

The transfer function of the Wiener filter is

$$\underline{w}^*(z) \triangleq \sum_{k=-\infty}^{\infty} \underline{w}^*(k) z^{-k} \tag{2.10}$$

For specific values of an individual matrix the optimal Wiener transfer function can be written as

$$w^*(z) = \frac{S_{xd}(z)}{S_{xx}(z)} \tag{2.11}$$

Consider now a single channel adaptive noise canceller with a typical set of inputs shown in Figure 2.3. The primary input consists of a signal s(j) plus the sum of two noises m0(j) and n(j). The reference input consists of a sum of two other noises m1(j) and n(j) * h(j), where h(j) is the impule response of the reference channel whose transfer function is H(z). To simplify the notation the transfer

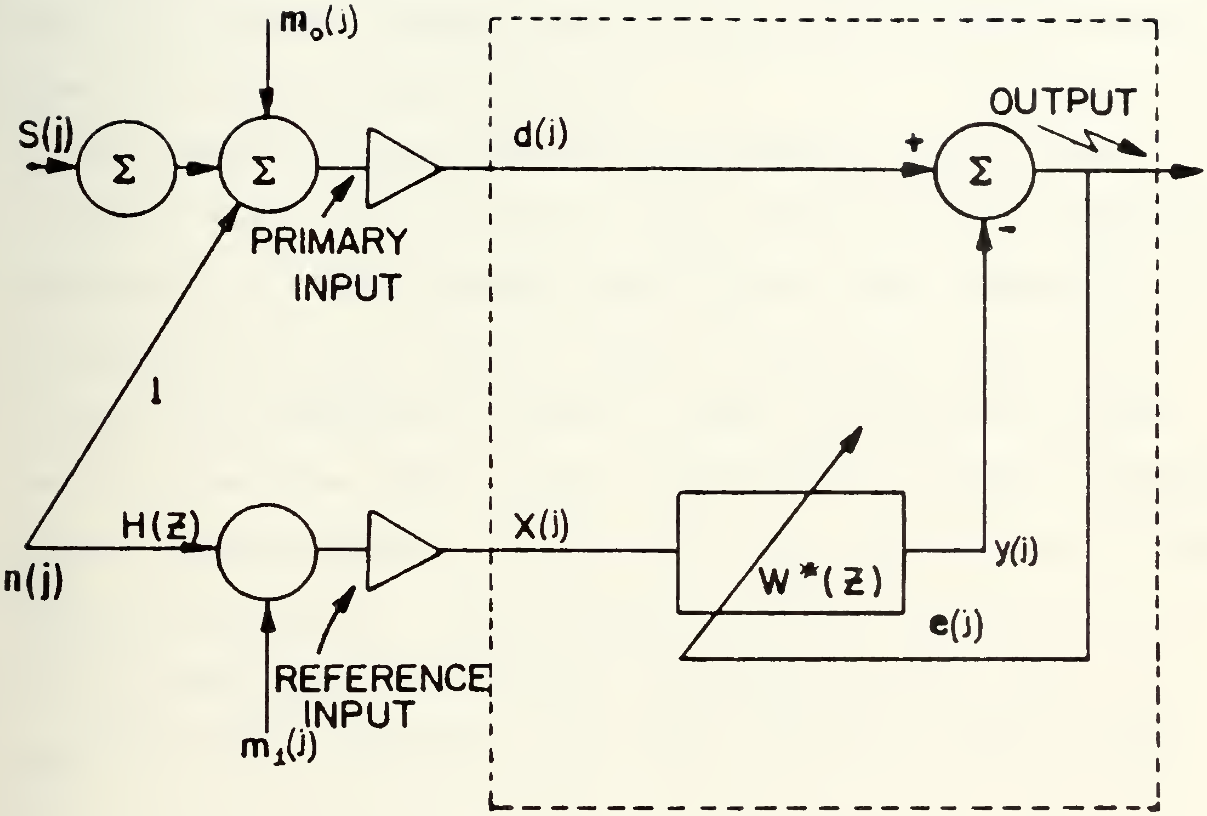

Fig. 2.3. Adaptive Noise Canceller with Correlated and Uncorrelated Noises in the Primary and Reference Inputs

function of the noise path from n(j) to the primary input has been set at unity. This procedure does not restrict the analysis since a suitable choice of H(z) and of statistics for n(j) will allow any combination of mutually correlated noises to appear at the primary and reference inputs. The noises m0(j) and m1(j) are uncorrelated with each other, with s(j) and with n(j) and n(j) * h(j). The noises n(j) and n(j) * h(j) have a common origin and are correlated with each other but uncorrelated with s(j).

The noise canceller includes an adaptive filter whose reference input x(j) is m1(j) + n(j) * h(j) and whose desired response d(j) is the primary input to the noise canceller and is composed of s(j) + m0(j) + n(j). The error signal e(j) is the noise canceller's output. When the adaptive process has converged then the optimal unconstrained transfer function of the adaptive filter is given by (2.11), which can be further reduced as follows. The spectrum of the noise m1 is $S_{m1m1}(z)$ and that of the noise n arriving via H(z) is $S_{nn}(z)|H(z)|^2$. Therefore, the input spectrum to the filter is

$$S_{xx}(z) = S_{m1m1}(z) + S_{nn}(z)|H(z)|^2 \tag{2.12}$$

The cross power spectrum between the filter's input and the desired response depends only on the mutually correlated primary and reference inputs and is given by

$$S_{xd}(z) = S_{nn}(z)H(z^{-1}) \tag{2.13}$$

Substituting Equations (2.12) and (2.13) into (2.11) results in the Wiener transfer function and is given by

$$w^*(z) = \frac{S_{nn}(z)H(z^{-1})}{S_{m1m1}(z) + S_{nn}(z)\,|H(z)|^2} \tag{2.14}$$

The result is a transfer function which is independent of the primary signal spectrum $S_{ss}(z)$ and of the primary uncorrelated noise spectrum $S_{m0m0}(z)$.

An interesting special case which clearly brings out the function of the adaptive noise canceller is when the additive noise m1 in the reference input is zero. Then $S_{m1m1}(z)$ is zero and the optimal transfer function becomes

$$w^*(z) = 1/H(z) \tag{2.15}$$

That is, the adaptive noise canceller causes the n(j) noise to be perfectly nulled at the noise canceller output. The primary uncorrelated noise m0(j) remains uncancelled.

Consider an analysis of the performance of the adaptive noise canceller in terms of the ratio of the signal power density at the output, $\rho_{out}(z)$, to the noise power density at the primary input $\rho_{pri}(z)$. This ratio can be written as

$$\frac{\rho_{out}(z)}{\rho_{pri}(z)} = \frac{S_{nn}(z) + S_{m0m0}(z)}{S_{output\ noise}(z)} = \frac{\text{primary noise pwr spectrum}}{\text{output noise pwr spectrum}} \tag{2.16}$$

with the signal power spectrum factored and cancelled out of the numerator. Figure 2.3 shows that the output noise spectrum consists of the sum of three components--one due to the propagation of m0(j) directly to the output, another due to the propagation of m1(j) to the output via the transfer function W(z), and another due to the propagation of n(j) to the output via the transfer function $1 - H(z)W(z)$. The output noise power spectrum can then be written as

$$S_{output\ noise}(z) = S_{m0m0}(z) + S_{m1m1}(z)\,|W^*(z)|^2 + S_{nn}(z)\,|1 - H(z)W^*(z)|^2 \qquad (2.17)$$

Now, if the ratios of the spectra of the uncorrelated noises to the spectra of the correlated noises (noise-to-noise density ratios) at the primary and reference inputs are defined as

$$A(z) = \frac{S_{m0m0}(z)}{S_{nn}(z)} \qquad (2.18)$$

$$B(z) = \frac{S_{m1m1}(z)}{S_{nn}(z)\,|H(z)|^2} \qquad (2.19)$$

then the transfer function of (2.14) can be written as

$$W^*(z) = 1/[H(z)(B(z)+1)] \qquad (2.20)$$

and the output noise power spectrum (2.17) can be written as

$$S_{\text{output noise}}(z) = S_{m0m0}(z) + \frac{S_{m1m1}(z)}{|H(z)|^2 (B(z)+1)^2}$$

$$+ S_{nn}(z) \left|1 - \frac{1}{B(z)+1}\right|^2$$

$$= S_{nn}(z) A(z) + S_{nn}(z) \frac{B(z)}{B(z)+1} \qquad (2.21)$$

and the ratio of the output to the primary input noise power spectra given in (2.16) is

$$\frac{\rho_{out}(z)}{\rho_{pri}(z)} = \frac{S_{nn}(z)[1 + A(z)]}{S_{\text{output noise}}(z)}$$

$$= \frac{[1 + A(z)]}{A(z) + (B(z))/(B(z)+1)}$$

$$= \frac{[A(z) + 1](B(z) + 1]}{A(z) + A(z)B(z) + B(z)} \qquad (2.22)$$

This expression allows an estimation of the level of noise reduction to be expected with an ideal noise cancelling system. In such a system the signal propagates to the output with a transfer function of unity. From (2.22) it can be seen that the ability of a noise cancelling system to reduce noise is limited by the uncorrelated-to-correlated noise density ratios at the primary and reference inputs. The smaller in magnitude are A(z) and B(z), the greater will be $\rho_{out}(z)/\rho_{pri}(z)$ and the more effective the action of the canceller. The desirability of low levels of uncorrelated noise in both inputs is made even more evident

by considering the approximations

1) small A(z) $$\frac{\rho_{out}(z)}{\rho_{pri}(z)} \cong \frac{1 + B(z)}{B(z)} \tag{2.23}$$

2) small B(z) $$\frac{\rho_{out}(z)}{\rho_{pri}(z)} \cong \frac{1 + A(z)}{A(z)} \tag{2.24}$$

3) small A(z) and B(z) $$\frac{\rho_{out}(z)}{\rho_{pri}(z)} \cong \frac{1}{A(z) + B(z)} \tag{2.25}$$

Infinite improvement is implied by these relationships when both A(z) and B(z) are zero resulting in complete removal of noise at the system output and perfect signal reproduction. When both A(z) and B(z) are small other factors such as misadjustment caused by gradient estimation noise in the adaptive process and the finite length of the adaptive filter limit system performance. These factors are discussed at some length in [Ref. 3].

C. SIGNAL PROPAGATION IN THE REFERENCE INPUT

If it is reasonable to consider reference noise propagation into the primary signal input it is also reasonable to consider certain instances when the signal propagates to the reference input. The system depicting the adaptive noise canceller with signal components in the reference input is shown in Figure 2.4. The derivation which discusses how much of the signal is cancelled when a portion of the signal input

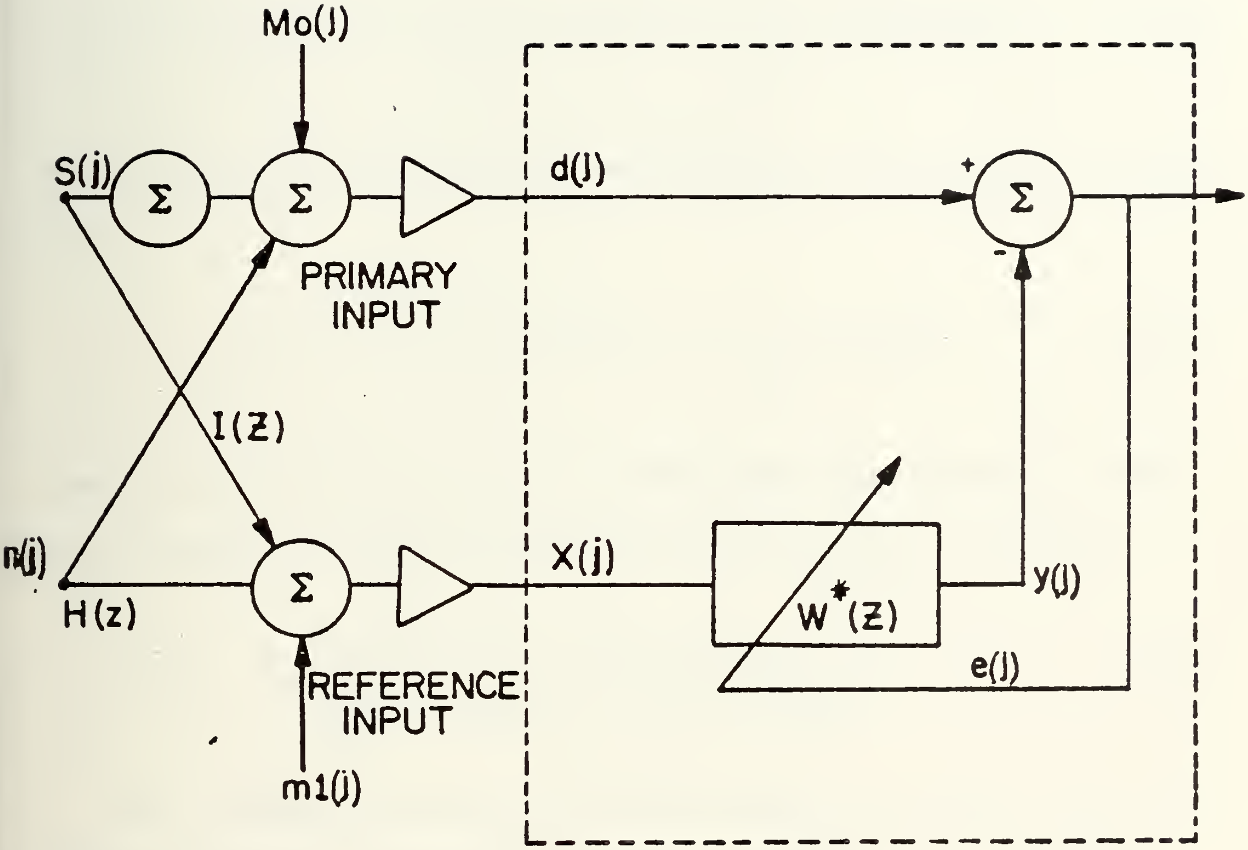

Fig. 2.4. Adaptive Noise Canceller With Signal Components in the Reference Input

"leaks" into the reference input omits the additive uncorrelated noises m0(j) and m1(j) in order to simplify the analysis.

Given that the spectrum of the signal is $S_{ss}(z)$ and that of the noise is $S_{nn}(z)$, then the spectrum of the reference input is given by

$$S_{xx}(z) = S_{ss}(z)|I(z)|^2 + S_{nn}(z)|H(z)|^2 \tag{2.26}$$

The cross spectrum between the reference and primary inputs is

$$S_{xd}(z) = S_{ss}(z)I(z^{-1}) + S_{nn}(z)H(z^{-1}) \tag{2.27}$$

When the adaptive process has converged, the Wiener transfer function of the adaptive filter given by (2.11) is

$$W^*(z) = \frac{S_{ss}(z)I(z^{-1}) + S_{nn}(z)H(z^{-1})}{S_{ss}(z)|I(z)|^2 + S_{nn}(z)|H(z)|^2} \tag{2.28}$$

The transfer function of the propagation path from the signal input to the noise canceller output is $1-I(z)W^*(z)$ and that of the path from the noise input to the canceller output is $1-H(z)W^*(z)$. The spectrum of the signal component in the output is thus

$$\begin{aligned} S_{ss\ out} &= S_{ss}(z)|1 - I(z)W^*(z)|^2 \\ &= S_{ss}(z)\left|\frac{|H(z)-I(z)|S_{nn}(z)H(z^{-1})}{S_{ss}(z)|I(z)|^2+S_{nn}(z)|H(z)|^2}\right|^2 \end{aligned} \tag{2.29}$$

and likewise, that of the noise is

$$S_{nn\ out}(z) = S_{nn}(z)\left|1 - H(z)W^*(z)\right|^2$$

$$= S_{nn}(z)\left|\frac{[I(z)-H(z)]S_{ss}(z)I(z^{-1})}{S_{ss}(z)|I(z)|^2 + S_{nn}(z)(H(z))^2}\right|^2 \qquad (2.30)$$

The output signal-to-noise density ratio is therefore

$$\rho_{out}(z) = \frac{S_{ss}(z)}{S_{nn}(z)}\left|\frac{S_{nn}(z)\ H(z^{-1})}{S_{ss}(z)\ I(z^{-1})}\right|^2$$

$$= \frac{S_{nn}(z)|H(z)|^2}{S_{ss}(z)|I(z)|^2} \qquad (2.31)$$

The output signal-to-noise density ratio can be conveniently expressed in terms of the signal-to-noise density ratio at the reference input in the following manner. The spectrum of the signal component in the reference input is

$$S_{ss\ ref} = S_{ss}(z)|I(z)|^2 \qquad (2.32)$$

and that of the noise component is likewise

$$S_{nn\ ref}(z) = S_{nn}(z)|H(z)|^2 \qquad (2.33)$$

Therefore, the signal-to-noise density ratio at the reference input is thus

$$\rho_{ref} = \frac{S_{ss}(z)\,|I(z)|^2}{S_{nn}(z)\,|H(z)|^2} \tag{2.34}$$

Comparison of Equation (2.34) with (2.31) shows that

$$\rho_{out}(z) = 1/\rho_{ref}(z) \tag{2.35}$$

This shows that if the noises in the primary and reference inputs are mutually correlated, the signal-to-noise density ratio at the noise canceller output is simply the reciprocal at all frequencies of the signal-to-noise ratio at the reference input. That is, in order to obtain a good signal-to-noise density ratio at the filter output there should be very little signal at the reference input.

The final objective of the analysis is to derive an expression for the spectrum of the output noise. As with the previous analysis it is instructive to first write the transfer function for the path from which the noise n(j) propagates to the output.

$$1 - H(z)W^*(z) = 1 - H(z)\,\frac{S_{ss}(z)I(z^{-1})+S_{nn}(z)H(z^{-1})}{S_{ss}(z)\,|I(z)|^2+S_{nn}(z)\,|H(z)|^2} \qquad 2$$

$$= \frac{S_{ss}(z)I(z^{-1})\,[I(z)-H(z)]}{S_{ss}(z)\,|I(z)|^2+S_{nn}(z)\,|H(z)|^2} \tag{2.36}$$

When I(z) is small (2.36) reduces to

$$1 - H(z)W^*(z) \;\tilde{=}\; \frac{-S_{ss}(z)I(z^{-1})}{S_{nn}(z)H(z^{-1})} \tag{2.37}$$

The output noise spectrum is

$$S_{\text{output noise}} = S_{nn}(z)\left|1 - H(z)W^*(z)\right|^2 \tag{2.38}$$

Again, considering the case where I(z) is small results in

$$S_{\text{output noise}}(z) \;\tilde{=}\; S_{nn}(z)\left|\frac{S_{ss}(z)I(z^{-1})}{S_{nn}(z)H(z^{-1})}\right|^2 \tag{2.39}$$

If Equation (2.39) is written in terms of the signal-to-noise density ratios at the reference and primary inputs, where the signal-to-noise density ratio at the reference input is given by Equation (2.34) and

$$\rho_{pri}(z) \triangleq \frac{S_{ss}(z)}{S_{nn}(z)} \tag{2.40}$$

then Equation (2.39) can be written as

$$S_{\text{output noise}} \;\tilde{=}\; S_{nn}(z)\left|\rho_{ref}(z)\right|\left|\rho_{pri}(z)\right| \tag{2.41}$$

Equation (2.41) shows that the output noise spectrum acts according to three factors (given that I(z) is small). First, the output noise spectrum depends on the input noise spectrum. Second, if the signal-to-noise density ratio at the reference input is low, the output noise will be low;

that is, the smaller the signal component feeding into the reference input, the better the cancellation of the noise. This is to be expected and was already shown by Equation (2.35). The third factor implies that if the signal-to-noise density ratio in the primary input (the desired response of the adaptive filter) is low, the filter will be trained most effectively to cancel the noise rather than the signal and therefore the output noise will be low.

As an illustration of the level of performance attainable in practical situations consider the following example. An adaptive noise cancelling system is designed to pass a plane-wave signal received in the main beam of an antenna array and to reject strong interference in the near field or in a minor lobe of the array. Assume that the signal and interference power spectra are overlapping and that the interference power density is twenty times greater than the signal power density at the individual array elements. Then the signal-to-noise ratio at the reference input ρ_{ref} is 1/20. Assume also that because of array gain the signal power equals the interference power at the array output which forms the primary input to the adaptive system. The signal-to-noise ratio at the primary input is $\rho_{pri} = 1$. After convergence the signal-to-noise ratio at the system output will be

$$\rho_{out} = 1/\rho_{ref} = 1/\frac{1}{20} = 20$$

If signal distortion is defined as

$$D(z) = \frac{\rho_{ref}(z)}{\rho_{pri}(z)} \tag{2.42}$$

then the maximum signal distortion will be

$$D(z) = (1/20)/1 = 5 \text{ percent}$$

The adaptive cancelling improves the signal-to-noise ratio twentyfold while introducing only 5 percent distortion. Additionally, the adaptive filter will provide the same performance when the input conditions change and a new set of convergent weights have been obtained.

D. NOISE CANCELLING APPLICATIONS

This section describes several practical applications which demonstrate the applied concepts of adaptive noise cancelling. These applications include cancellation of noise in speech signals, cancellation of antenna sidelobe interference, cancellation of 60-Hz interference and cancellation of either periodic or broadband interference when no reference is available.

A prime example of noise contaminated speech is that of a pilot communicating by radio from the cockpit of an aircraft where a high level of engine noise is interfering with the pilot's voice. The noise contains, among other components, strong periodic mixtures that occupy the same frequency band as speech. These components cannot be "low filtered" or "high filtered" out of the speech pattern and are picked up by

the microphone into which the pilot speaks, severely interfering with the intelligibility of the radio transmission. It is impractical to process the transmission with a conventional fixed filter because the frequency and intensity of the noise components vary with engine speed and load and even the position of the pilot's microphone. By placing a second microphone at a suitable location in the cockpit, a sample of the ambient noise which is free of the pilots speech can be obtained. This sample can be filtered and subtracted from the transmission, significantly reducing the interference.

Widrow et al., [Ref. 4] demonstrated the feasibility of cancelling noise in speech signals by conducting a number of experiments which simulated the cockpit noise problem. Figure 2.5 shows the system used for cancelling the noise in the cockpit noise simulations. A voice input from a room with strong acoustic interference was used as the primary input. A second microphone was placed in the room away from the speaker and this was used as the reference input. The output of the noise cancelling was then monitored by a remote listener. The canceller included an adaptive filter with 16 weights whose values were digitally controlled by a computer. A typical experiment used an audio frequency triangular wave containing many harmonics as interference. Because of multipath effects the amplitude and phase of the interference varied from point to point in the room. The noise cancelling system was able to reduce the output power of the interference, which otherwise made the speech unintelligible, by 20 to 25

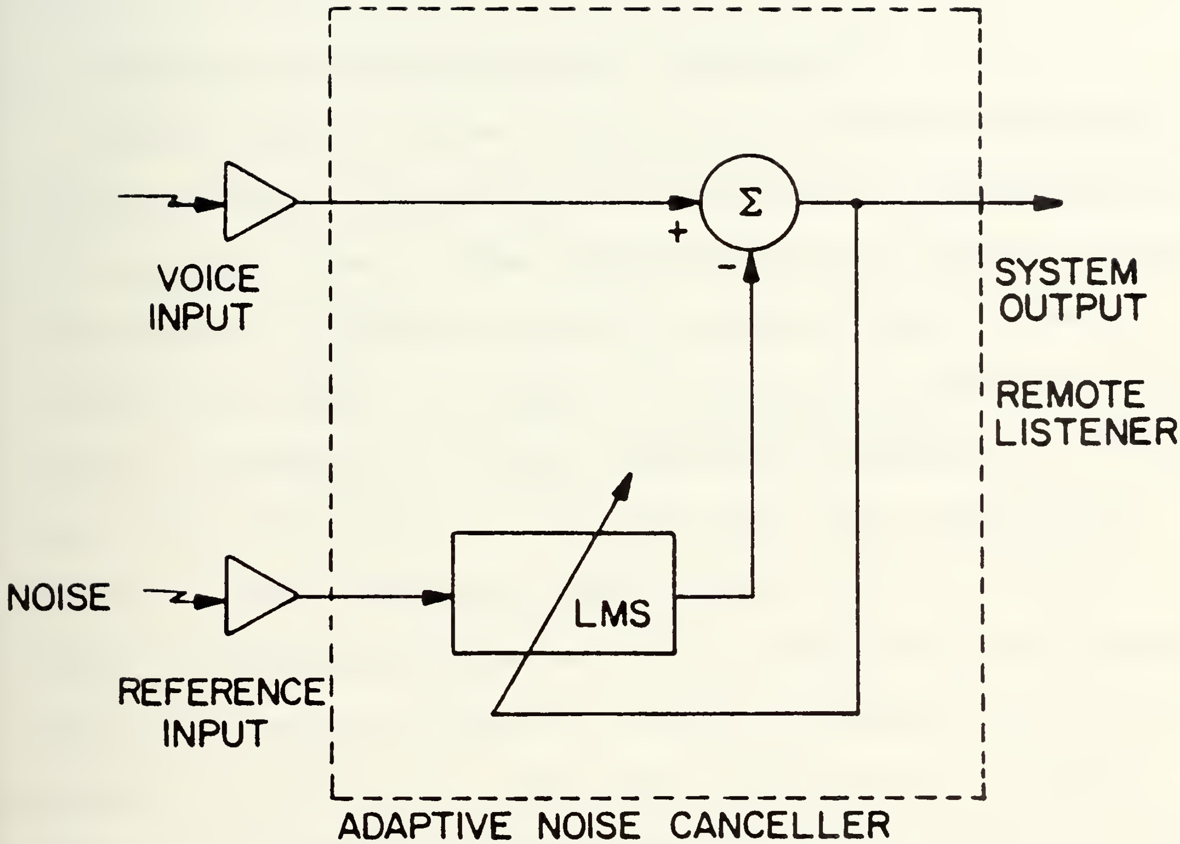

Fig. 2.5. Cancelling Noise in Speech Signals

dB, rendering the interference barely perceptible to the remote listener. No noticeable distortion was introduced into the speech pattern. Convergence times were on the order of a few seconds and the system was readily able to readapt when the position of one or both microphones was changed or when the frequency of the interference was varied over the range of 100 to 2000 Hz.

E. CANCELLING ANTENNA SIDELOBE INTERFERENCE

Another type of noise cancelling is that of eliminating strong unwanted signals which are incident on the sidelobes of an antenna array. These interferences can severely retard the reception of weaker signals on the main beam. The conventional method of reducing this type of interference by adaptive beamforming is often complex and expensive to implement. When the number of spatially discrete interference sources is small, adaptive noise cancelling can provide a simpler and less expensive method of coping with this problem. Consider an array pattern with signal strengths and directions as shown in Figure 2.6. The array consists of a circular pattern of 16 equally spaced omnidirectional elements. The outputs of the elements are delayed and summed to form a main beam steered at a relative angle of 0 degrees. A simulated "white" signal consisting of uncorrelated samples of unit power is assumed to be incident on the beam. Simulated interference with the same bandwidth and with a power of 100 is incident on the main beam at a realtive angle of 58 degrees.

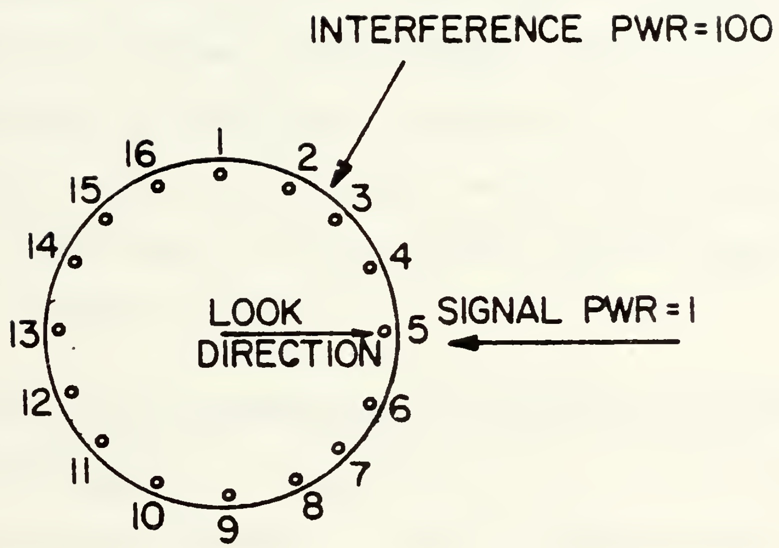

Fig. 2.6. Array Configuration for Adaptive Sidelobe Canceller

The entire array is then connected to the adaptive noise cancelling system shown in Figure 2.7. In this case the output of the beamformer serves as the canceller's primary input, and the output of one of the elements (#4) is arbitrarily chosen as the reference input. The adaptive canceller uses 14 weights.

A number of experiments performed in [Ref. 4] show that the signal-to-noise ratio at the system output was found after convergence to be +20 dB. The signal-to-noise ratio at the single array element was -20 dB. This result bears out the equation shown in the Wiener solution (2.35), that the signal-to-noise ratio at the system output would be the reciprocal of the ratio at the reference input, which is derived from a single element.

F. CANCELLING 60-HZ INTERFERENCE IN ELECTROCARDIOGRAPHY

A practical example of cancelling 60-Hz interference is found in electrocardiography. A major problem which exists in the recording of electrocardiograms (ECG's) is the appearance of unwanted 60-Hz interference in the output. Various methods have been utilized to help cancel the 60-Hz interference, including more effective grounding techniques and the use of twisted cabling. Another method capable of reducing 60-Hz ECG interference is adaptive noise cancelling via a system such as that shown in Figure 2.8.

The primary input is taken from the ECG preamplifier and the 60-Hz reference is taken from a wall outlet. The adaptive

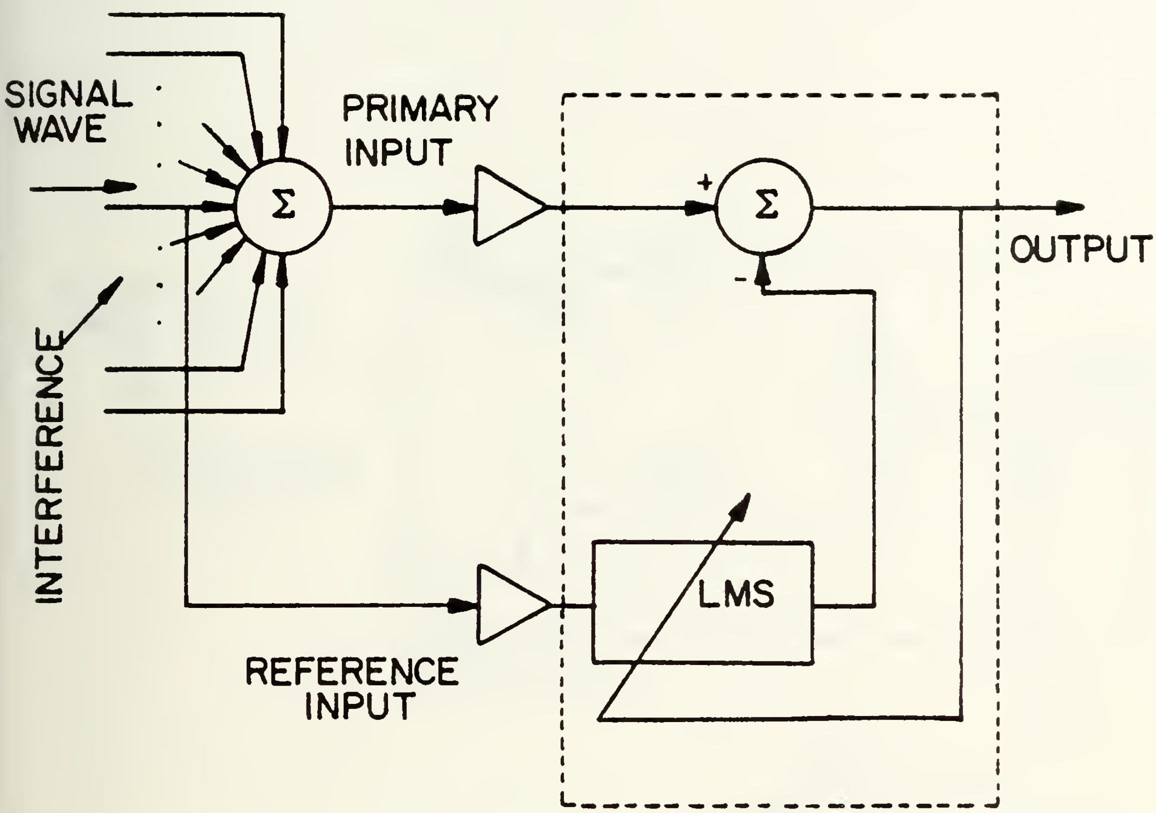

Fig. 2.7. Adaptive Noise Cancelling Applied to a Receiving Array

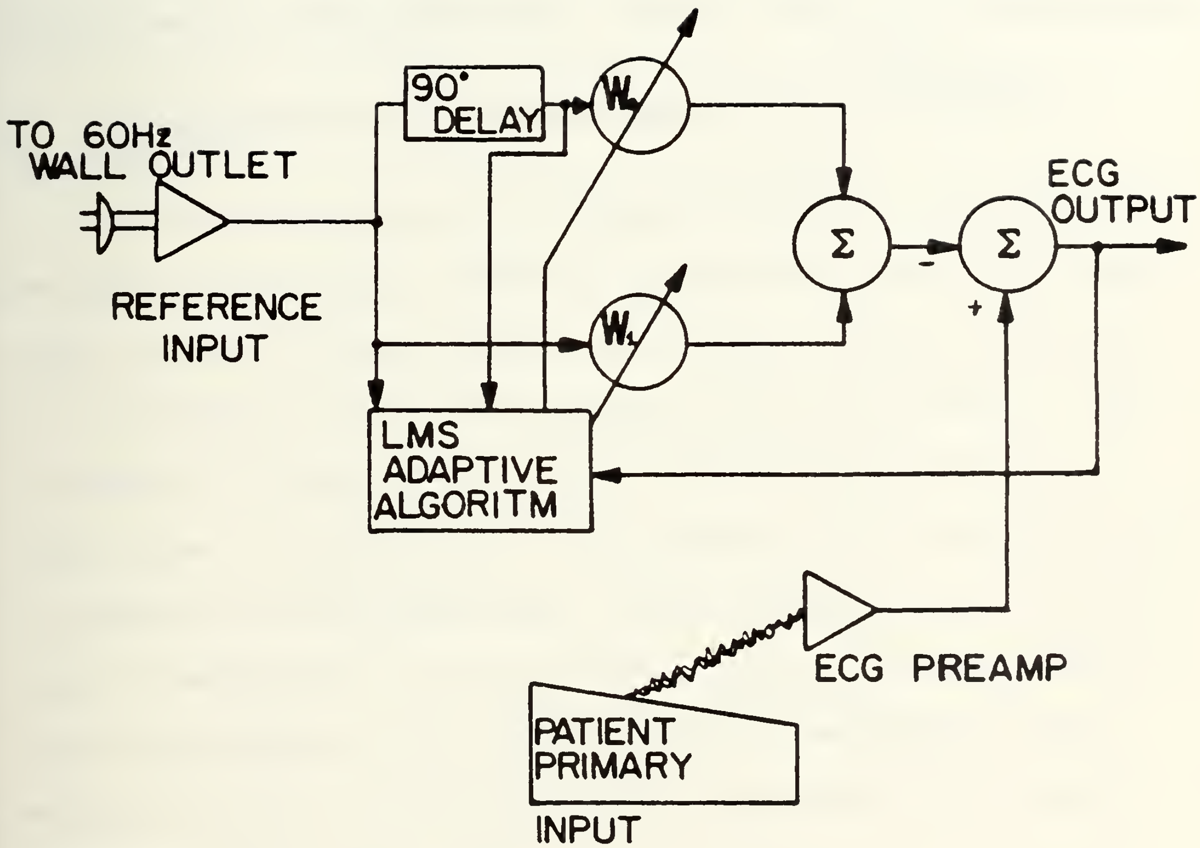

Fig. 2.8. Cancelling 60-Hz Interference in Electrocardiography

filter is simple, containing only two variable weights, one applied to the reference input directly and the other to a version of it shifted in phase by 90 degrees. The two weighted versions of the reference are summed to form the filter's output, which is then subtracted from the primary input. A valuable advantage in the use of an adaptive filter rather than a fixed notch filter at 60-Hz, is that the variable weights allow the 60-Hz interference to change in both magnitude and phase and still realize effective cancellation.

G. CANCELLING PERIODIC INTERFERENCE WITH NO EXTERNAL REFERENCE

In many cases where a broadband signal is corrupted by periodic interference there is no external reference input which is free of the signal. If a fixed delay is inserted in a reference input drawn directly from the primary input, as shown in Figure 2.9, the periodic interference can, in many cases, be cancelled. A key point is that the delay must be chosen to be of sufficient length to cause the broadband signal components in the reference input to become decorrelated from those in the primary input. Because of their periodic nature, the interference components will remain correlated with each other.

By taking the output from the LMS output instead of the difference signal the same system can be used to separate broadband interference from a periodic signal as shown in Figure 2.10. Such a filter is often called an adaptive

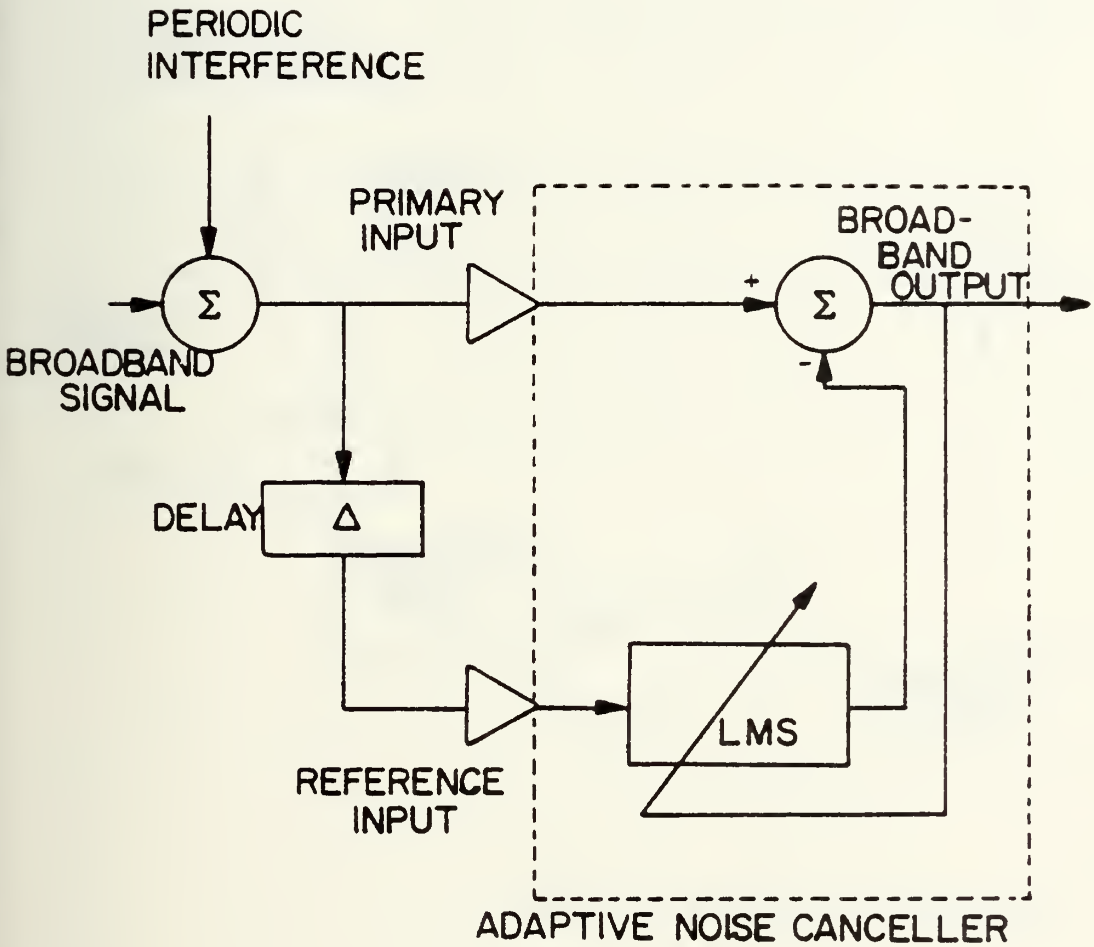

Fig. 2.9. Adaptive System for Cancelling Periodic Interference Without an External Reference

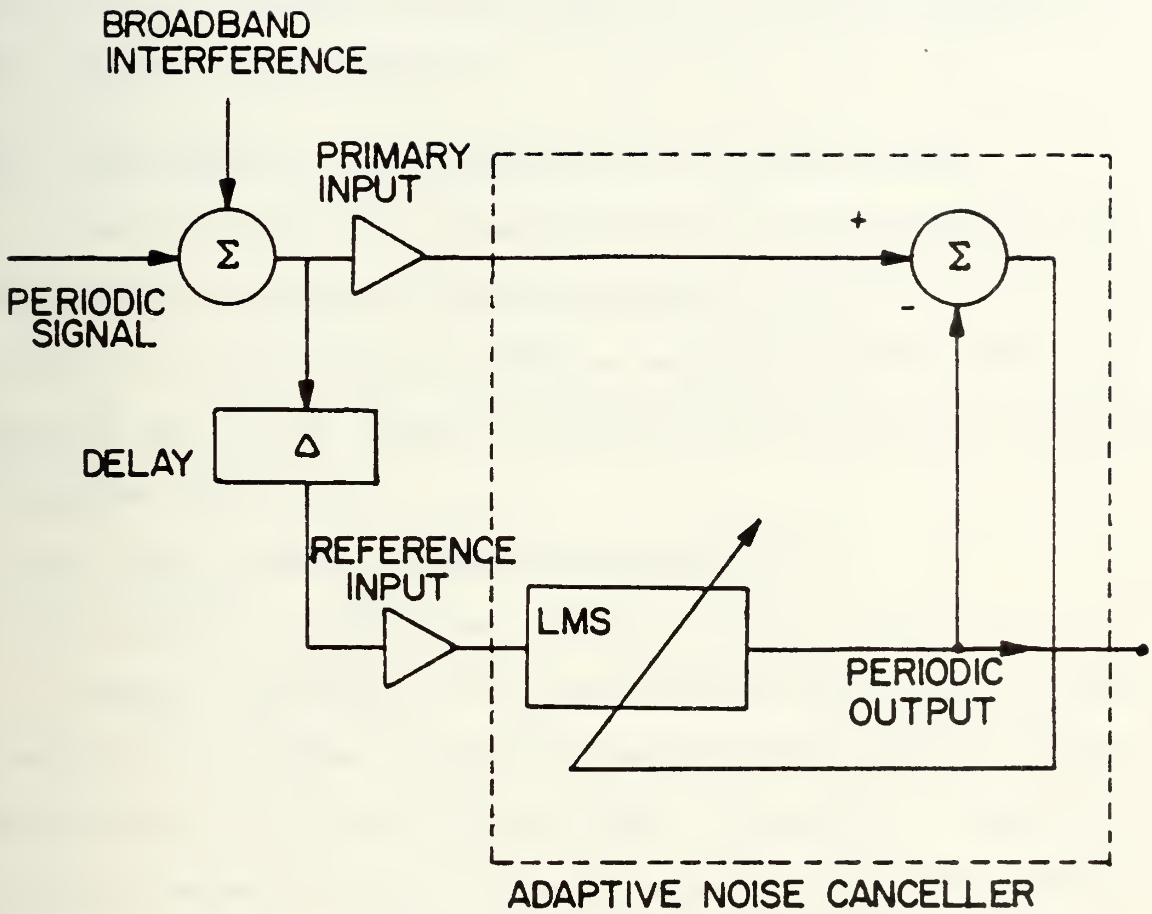

Fig. 2.10. Adaptive System for Cancelling Broadband Interference Without an External Reference

self-tuning filter. Additionally, applications of this filter are utilized in the adaptive line enhancer, a system used for detection of a low level signal imbeded in noise. The transfer function of this filter is the digital Fourier transform of the impulse response. Its magnitude at the frequency of the interference is very nearly one, the value required for perfect cancellation.

H. THE ADAPTIVE NOISE CANCELLER AS A NOTCH FILTER

One of the primary considerations for considering constraints on adaptive filtering systems stems from the filter's use as a notch filter. After an analysis of the adaptive system as a notch filter it was felt that specific constraints could be used to

a) shape the frequency characteristics

b) produce a faster convergence.

Before considering the constraining equations and their results it is first helpful to analyze the adaptive noise canceller in its notch filter mode. Figure 2.11 depicts a single frequency adaptive noise canceller with two adaptive weights. Analytical and experiemntal results show that if more than one frequency is present in the reference input then a notch for each will be formed. The primary input is assumed to be any type of signal--stochastic, deterministic, periodic, transient, or any combination of these. The reference input is assumed to be a pure cosine wave $C \cos(\omega_0 t + \phi)$. The primary and reference inputs are sampled

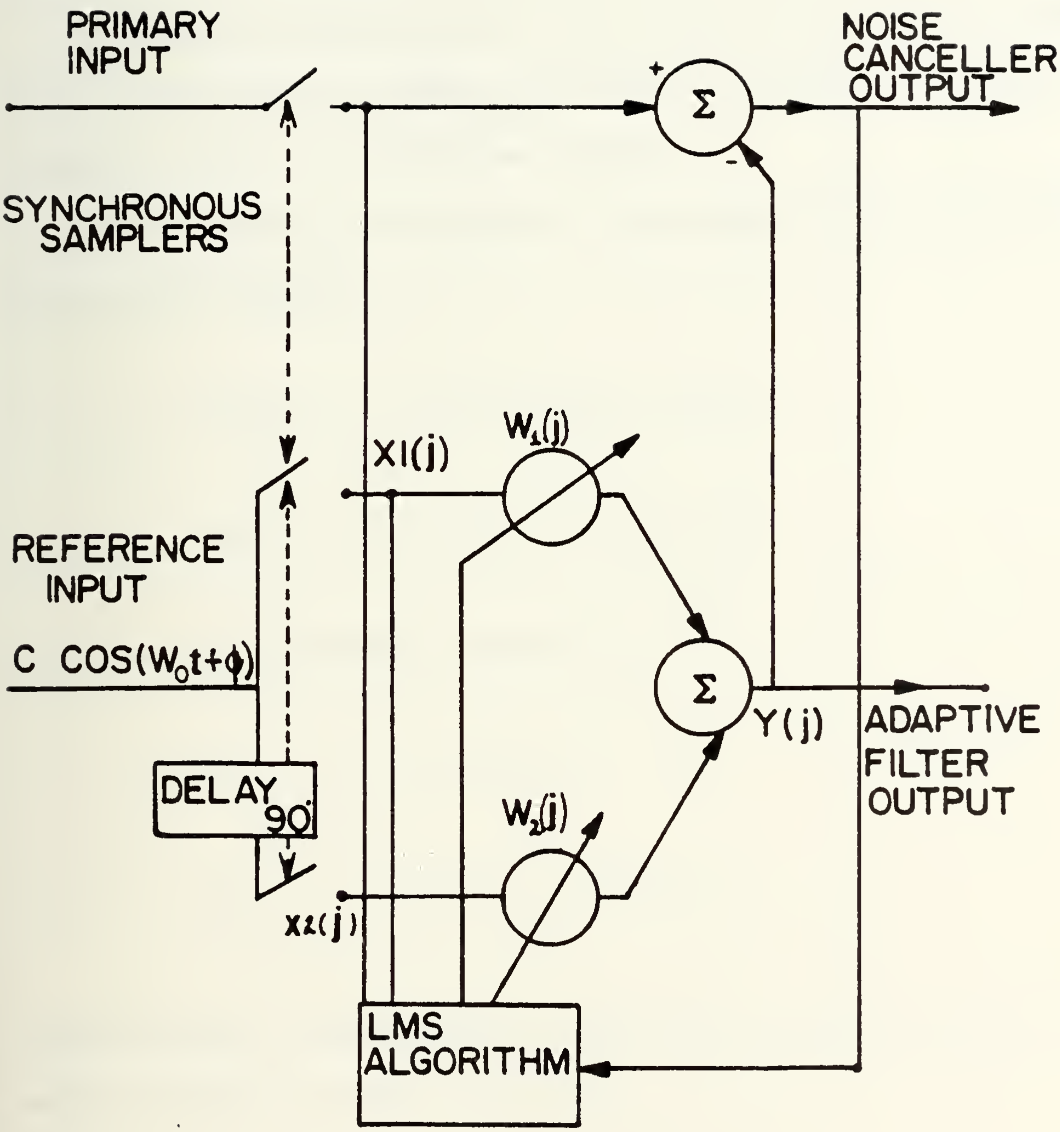

Fig. 2.11. Single Frequency Adaptive Noise Canceller

at the frequency of $\omega_s = 2(\text{Pi})/T$ rad/sec. The reference input is sampled directly yielding X1(j), and after undergoing a 90 degree phase shift, also produces X2(j). Assume synchronous sampling.

A transfer function for the noise canceller of Figure 2.8-1 can be obtained by analyzing signal propagation from the primary input to the system output.

The weights are updated in accordance with the LMS algorithm yielding

$$
\begin{aligned}
W1(j+1) &= W1(j) + 2\mu\varepsilon(j)X1(j) \\
W2(j+1) &= W2(j) + 2\mu\varepsilon(j)X2(j)
\end{aligned}
\tag{2.43}
$$

The sampled reference inputs are

$$
\begin{aligned}
X1(j) &= C\cos(\omega_o(j)T + \phi) \\
X2(j) &= C\sin(\omega_0(j)T + \phi)
\end{aligned}
\tag{2.44}
$$

Using signal flow diagram techniques and considering that the error signal at time j = k is

$$\varepsilon(j) = \delta(j-k) \tag{2.45}$$

the filter's impulse response at k = 0 can be given as

$$y(j) = 2\mu C^2 u(j-1)\cos(\omega_0(j)T) \tag{2.46}$$

where u(j) is the discrete unit step function. The transfer function of this path is

$$G(z) = 2\mu C^2 \frac{z(z - \cos\omega_0 T)}{z^2 - 2z\cos\omega_0 T + 1} - 1$$

$$= \frac{2\mu C^2 (z \cos \omega_0 T - 1)}{z^2 - 2z \cos \omega_0 T + 1} \qquad (2.47)$$

This transfer function can be expressed in terms of the radian sampling frequency $\omega_s = 2(\text{Pi})/T$ as

$$G(z) = \frac{2\mu C^2 [z \cos (2\pi\omega_0\omega_s^{-1}) - 1]}{z^2 - 2z \cos (2\pi\omega_0\omega_s^{-1}) + 1} \qquad (2.48)$$

When the feedback loop from the adaptive filter output to the difference signal is formed, the transfer function H(z) from primary input to noise canceller output can be written as

$$H(z) = \frac{z^2 - 2z \cos (2\pi\omega_0\omega_s^{-1}) + 1}{z^2 - 2(1 - \mu C^2) z \cos (2\pi\omega_0\omega_s^{-1}) + 1 - 2\mu C^2} \qquad (2.49)$$

This transfer function has the property of a notch filter at the reference frequency ω_0. The zeros of the transfer function are located in the z-plane at

$$z = \exp(\pm j 2\pi\omega_0\omega_s^{-1}) \qquad (2.50)$$

and are exactly on the unit circle at angles of $\pm 2\pi\omega_0\omega_s^{-1}$ radians. The poles are inside the unit circle at a radial

distance $(1-2\mu C^2)^{1/2}$ from the origin. For slow adaptation (small values of μC^2) the angles of the poles are almost identical to the zeros. Since the zeros lie on the unit circle, the depth of the notch in the transfer function is infinite at frequency $\omega = \omega_0$. The sharpness of the notch is determined by the closeness of the poles to the zeros. Corresponding poles and zeros are separated by a distance approximately equal to μC^2. The notch bandwidth is given by

$$BW = \mu C^2 \omega_s / \pi \tag{2.51}$$

and the Q of the notch is determined by the ratio of the center frequency to the bandwidth.

$$Q \cong \frac{\omega_0}{\mu C^2 \omega_s} \tag{2.52}$$

The single frequency noise canceller is, therefore, equivalent to a stable notch filter when the reference is a pure cosine wave. The depth of the null is generally superior to that of a fixed digital or analog filter because the adaptive process maintains the null exactly at the reference frequency.

III. THE CONSTRAINED ADAPTIVE FIR FILTER

In this chapter the concept of the constrained adaptive filter is introduced. A literature search has indicated that very little research has been done in the area of constrained adaptive filters. Frost [Ref. 8] presents a constrained LMS algorithm which is capable of adjusting an array of sensors in real time to respond to a signal coming from a desired direction while discriminating against noises coming from other directions. A set of linear constraints on the weights maintains a chosen frequency characteristic for the array in the direction of interest.

In this chapter three constraint conditions are presented. The first two involve a constraint on the angle between zero locations, so that this angle remains constant while the zero location changes adaptively. The constraint essentially maintains the pattern of the zeros while their location is shifted. The first approach is a direct implementation in which one of the weights is changed adaptively using the LMS algorithm and the others are slaved by the fixed angle formula to the adaptively adjusted weight. The second implementation involves a cascaded version of the foregoing master-slave concept. The third constraint considered is a linear constraint on the weights. The approach presented uses a Lagrangian multiplier formulation to augment the cost function in which the basic LMS adaptive algorithm is applied.

A. FIXED ANGLE CONSTRAINT--DIRECT IMPLEMENTATION

Consider a transversal filter using the LMS algorithm with the constraint that the angular separation between the filter zeros in the z-plane is to be a constant as set by design requirements. As an example, consider a fourth order FIR filter with $B = \theta_2 - \theta_1$, where θ_1 and θ_2 are the angles of the zeros. The value of B and the magnitude of the zeros is to be kept constant while the angles θ_1 and θ_2 are to be adjusted adaptively. This specification serves to maintain the shape of the filter characteristics in the frequency domain. From the direct implementation of the transversal filter shown in Figure 3.1 the system transfer function is given by

$$H(z^{-1}) = 1 + W_1 z^{-1} + W_2 z^{-2} + W_3 z^{-3} + W_4 z^{-4} \tag{3.1}$$

$$= (z^{-4})(z^4 + W_1 z^3 + W_2 z^2 + W_3 z + W_4) \tag{3.2}$$

which can be factored into the form

$$H(z) = \frac{N(z)}{D(z)} = \frac{(z^2 - 2r_1\cos\theta_1 z + r_1^2)(z^2 - 2r_2\cos\theta_2 z + r_2^2)}{z^4} \tag{3.3}$$

where r_1 and r_2 represent the radius of the zero placement in the z-plane and θ_1 and θ_2 are their respective angular displacement from the axis. Carrying out the indicated multiplication of the numerator of (3.3) yields

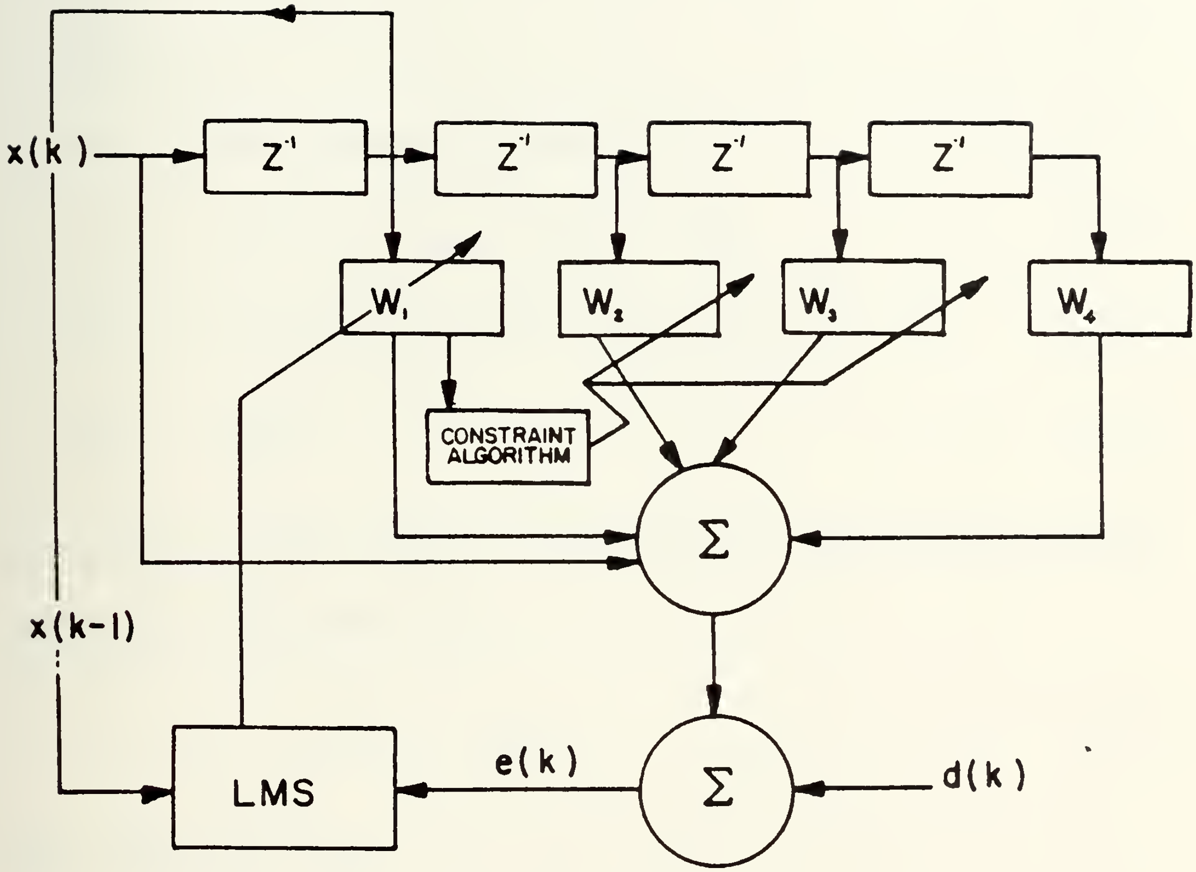

Fig. 3.1. Transversal Adaptive Filter - Direct Implementation

$$N(z) = z^4-(2r_1\cos\theta_1+2r_2\cos\theta_2)z^3 + (r_1^2+4r_1r_2\cos\theta_1\cos\theta_2+r_2^2)z^2$$

$$- (2r_1^2r_2\cos\theta_2 + 2r_2^2r_1\cos\theta_1)z + r_1^2r_2^2 \qquad (3.4)$$

Setting these terms equal to the weight values of (3.2) yields

$$W_0 = 1 \qquad (3.5)$$

$$W_1 = -2(r_1 \cos\theta_1 + r_2 \cos\theta_2) \qquad (3.6)$$

$$W_2 = (r_1^2 + 4r_1r_2 \cos\theta_1 \cos\theta_2 + r_2^2) \qquad (3.7)$$

$$W_3 = (-2r_1r_2)(r_1 \cos\theta_2 + r_2 \cos\theta_1) \qquad (3.8)$$

$$W_4 = r_1^2r_2^2 \qquad (3.9)$$

The adaptive LMS algorithm must now be constrained so that the angle B is given by

$$B \triangleq \theta_1 - \theta_2 \qquad (3.10)$$

For simplicity let

$$r_1 = r_2 = R \qquad (3.11)$$

From (3.10) it follows that

$$\theta_1(k+1) = B + \theta_2(k+1) \qquad (3.12)$$

Using (3.11) and (3.12) Equation (3.6) can now be written as

$$W_1(k+1) = -2R(\cos(B + \theta_2(k+1)) + \cos(\theta_2(k+1)) \qquad (3.13)$$

Using the trigonometric identity

$$\cos(A+B) = \cos A \cos B - \sin A \sin B$$

Equation (3.13) can be written as

$$W_1(k+1) = -2R(\cos B \cos(\theta_2(k+1)) - \sin B \sin(\theta_2(k+1))$$

$$+ \cos(\theta_2(k+1)) \qquad (3.14)$$

Combining terms, and defining the constants

$$K3 = -2R(\cos(B) + 1) \qquad (3.15)$$

$$K4 = -2R(\sin(B)) \qquad (3.16)$$

Equation (3.13) can finally be written as

$$W_1(k+1) = K3 \cos(\theta_2(k+1)) - K4 \sin(\theta_2(k+1)) \qquad (3.17)$$

This transcendental equation can be solved iteratively for $\theta_2(k+1)$ using the value for $W_1(k+1)$ obtained from the LMS algorithm

$$W_1(k+1) = W_1(k) + 2\mu X(k-1)\varepsilon(k) \qquad (3.18)$$

Now that $\theta_2(k+1)$ is known, Equation (3.7) can be solved for $W_2(k+1)$, where

$$W_2(k+1) = 2R^2 + 4R^2 \cos(\theta_2(k+1) + B) \cos(\theta_2(k+1)) \quad (3.19)$$

Recognizing that Equation (3.8) is proportional to Equation (3.6), the fourth algorithm iteration can be written as

$$W_3(k+1) = R^2 W_1(k+1) \quad (3.20)$$

From (3.9) W_4 is a constant value

$$W_4(k+1) = R^4 \quad (3.21)$$

The four equations (3.18) through (3.21) comprise an adaptive iterative algorithm that produces the direct analytical realization of Figure 3.1. It is seen that W_1 is changed adaptively and W_2 and W_3 are slaved to W_1, and W_0 and W_4 are constants. This approach yields two pairs of zeros which are located at radius R from the origin of the z plane and precisely B degrees apart. Figure 3.2 depicts the resulting z-plane diagram.

B. FIXED ANGLE CONSTRAINT--CASCADED IMPLEMENTATION

A cascaded implementation of the fourth order transversal filter is shown in Figure 3.3. The transfer functions resulting from Figure 3.3 are given by

$$H_1(z) = \frac{z^2 - 2r_1 \cos\theta_1 z + r_1^2}{z^2} \quad (3.22)$$

for the first section, and for the second section

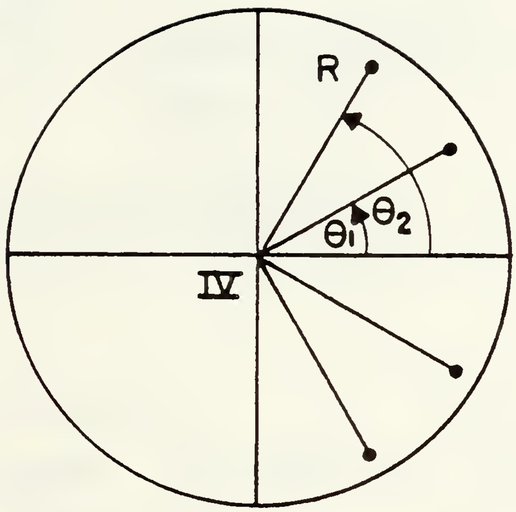

Fig. 3.2. Z-Plane Diagram of Zeros Resulting from Transversal Filter
IV indicates four poles located at the Origin

Fig. 3.2. S-Plane Diagram of Zeros Resulting fro
Transversal Filter
IV indicates four poles located at the
origin

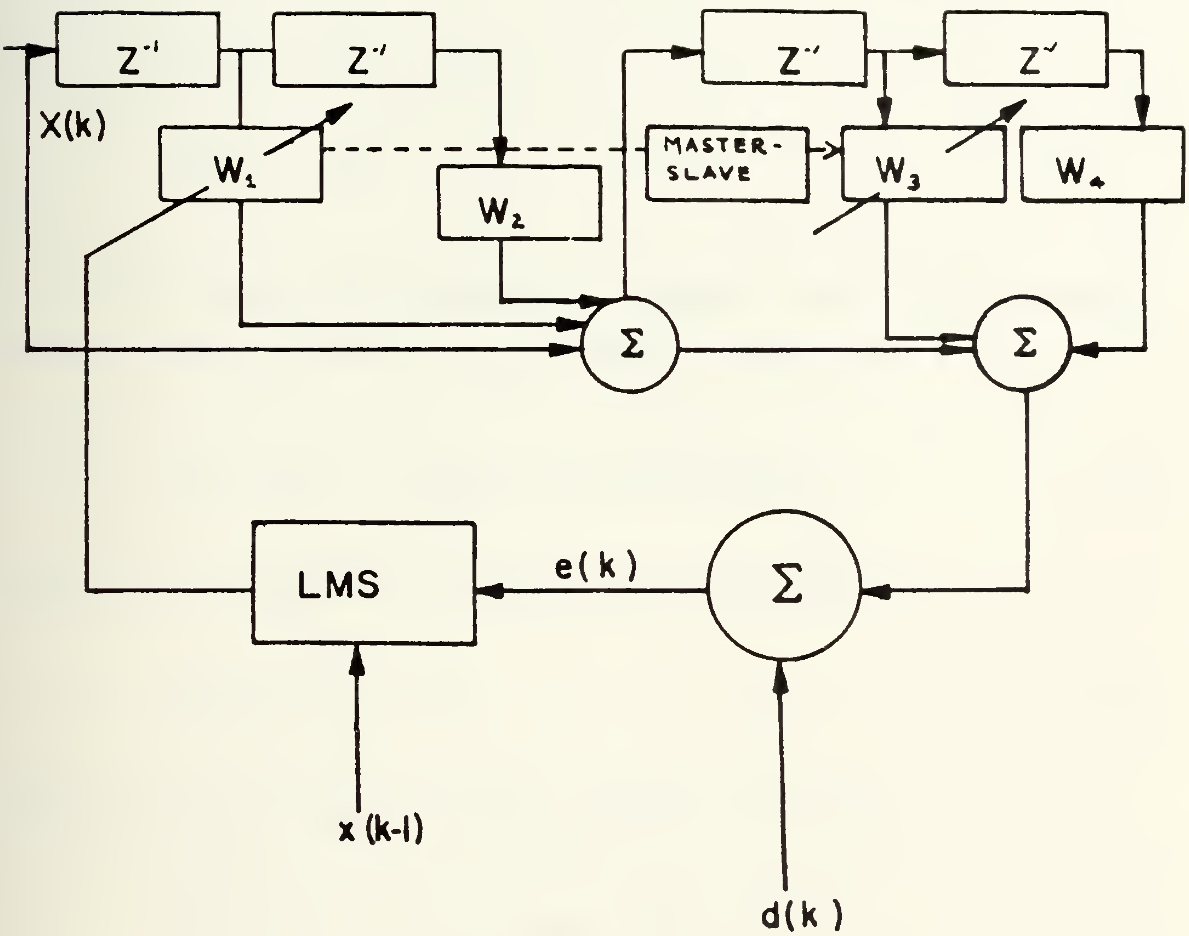

Fig. 3.3. Cascaded Transversal Filter with LMS Adaptation of W_1. Master-Slave Concept

$$H_2(z) = \frac{z^2 - 2r_2 \cos\theta_2 z + r_2^2}{z^2} \tag{3.23}$$

The weight values from these two equations are

$$W_1 = -2\, r_1 \cos\, \theta_1$$

$$W_2 = r_1^2$$

$$W_3 = -2\, r_2 \cos\, \theta_2$$

$$W_4 = r_2^2$$

Once again the algorithm producing a set of converging weights begins with the LMS adaptive equation

$$W_1(k+1) = W_1(k) + 2\mu X(k-1)\varepsilon(k) \tag{3.24}$$

The second weight is a constant proportional to

$$W_2(k+1) = R^2 \tag{3.25}$$

From (3.22) the value for W_1 can be written as

$$W_1(k+1) = -\, 2\, r_1 \cos\, \theta_1(k+1) \tag{3.26}$$

Solving for $\theta_1(k+1)$

$$\cos\theta_1(k+1) = \frac{-W_3(k+1)}{2r_1} \tag{3.27}$$

where the value of $W_1(k+1)$ is known from (3.24) and

$$r_1 = r_2 = R$$

Therefore,

$$\theta_1(k+1) = \text{Arccosine}\left(\frac{-W_1(k+1)}{2R}\right) \qquad (3.28)$$

Now the third weight can be written in the form of

$$W_3(k+1) = -2\,R\cos\theta_2(k+1) \qquad (3.29)$$

where

$$\theta_2(k+1) = B - \theta_1(k+1) \qquad (3.30)$$

with B a given constant set by the user and $\theta_1(k+1)$ known from (3.27). The fourth weight is proportional to

$$W_4(k+1) = R^2 \qquad (3.31)$$

The four equations (3.24), (3.25) and (3.29), and (3.31) comprise an adaptive iterative algorithm that produces the cascaded realization of Figure 3.3. As in the direct realization, the solution of Equations (3.22) and (3.23) using the adaptive weight solutions produces two pairs of zeros which are located at radius $r_1 = r_2 = R$ from the origin of the z-plane and precisely B degrees apart.

C. LINEARLY CONSTRAINED WEIGHTS--LAGRANGE MULTIPLIER APPROACH

Consider now the LMS adaptive equations modified for a different type of constraint, namely a linear constraint on

the sum of the weight vector. That is, the constraint that

$$W_1(k) + W_2(k) + \ldots + W_n(k) = K \tag{3.22}$$

where K is some user defined constant. This constraint function is adjoined to the square of the error equation by the method of Lagrange multipliers [Ref. 9]. Recalling from Equation (1.5) that

$$\varepsilon(k) = d(k) - \underline{W}(k)^T\underline{X}(k)$$

the adjoined and modified least mean squared error function can be written as the cost function equation

$$J(k) = [d(k)-\underline{W}(k)^T\underline{X}(k)]^2 - \lambda(k)[\underline{W}(k)^T\underline{u} - K] \tag{3.33}$$

where

$$\underline{u} = [1 \quad 1 \quad \ldots \quad 1]^T$$

and $\lambda(k)$ is the scalar Lagrange multiplier.

In order to minimize the mean squared error under the given constraint consider the gradient of (3.33) with respect to the weights and to the Lagrangian multipliers.

$$\nabla[J(k)]_{\underline{W}} = -2[d(k)-\underline{W}^T(k)\underline{X}(k)]\underline{X}(k) - \lambda(k)\underline{u} \tag{3.34}$$

and

$$\nabla[J(k)]_{\lambda} = -[\underline{W}(k)^T\underline{u} - K] \qquad (3.35)$$

The method of steepest descent can be described by the two relationships

$$\underline{W}(k+1) = \underline{W}(k) + k_1\nabla[J(k)]_{\underline{W}} \qquad (3.36)$$

and

$$\lambda(k+1) = \lambda(k) + k_2\nabla[J(k)]_{\lambda} \qquad (3.37)$$

Substituing Equations (3.34) and (3.35) into Equations (3.36) and (3.37) yields the final weight and Lagrange parameter iteration algorithms using the LMS approximation

$$\underline{W}(k+1) = \underline{W}(k) - 2k_1 e(k)\underline{X}(k) - \lambda(k)\underline{u} \qquad (3.38)$$

and

$$\lambda(k+1) = \lambda(k) - k_2[\underline{W}(k)^T\underline{u} - K] \qquad (3.39)$$

where

$\underline{W}(k)$ = the weight vector before adjustment

$\underline{W}(k+1)$ = the weight vector after adjustment

K = the linear constraint value

k_1, k_2 = the scalar constants ($k_1, k_2 < 0$)

$\lambda(k)$ = the Lagrange parameter before adjustment

$\lambda(k+1)$ = the Lagrange parameter after adjustment.

The resulting constrained adaptive filter is shown in Figure 3.4. Experimental results are presented in the next chapter.

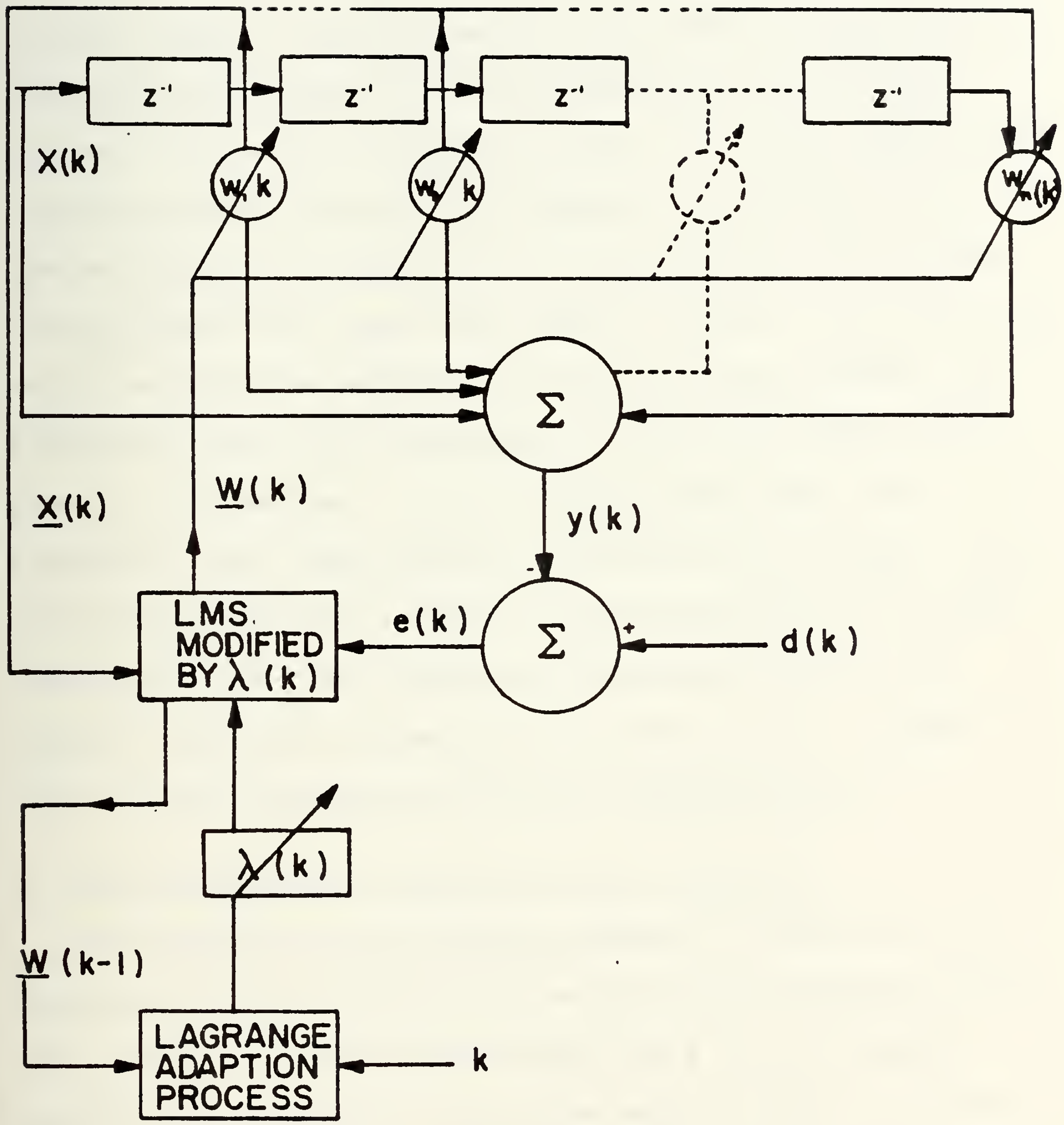

Fig. 3.4. The Adaptive LMS Filter with Linear Constraint and Lagrange Parameters

IV. SIMULATIONS AND RESULTS

Using the Hewlett-Packard 85 (HP-85) microcomputer several computer simulations are performed to demonstrate both the unconstrained and the constrained LMS adaptive systems discussed in Chapters II and III. The unconstrained adaptive system utilized for computer simulation is the system depicted in Figure 2.1. The unconstrained results are obtained first with nine adaptive weights and then with only four adaptive weights. The noise input consists of either the sum of two sinewaves or zero mean white noise added to a sinewave. In all unconstrained cases the reference input (the desired waveform) is a sinewave. The constrained adaptive systems are simulated for the linear constraint using the Lagrangian multiplier and for the angular constraint (separation of zeros in the z-plane plot) in both cascaded and direct implementations.

A. THE UNCONSTRAINED ADAPTIVE NOISE CANCELLER

The adaptive noise canceller shown in Figure 2.1 is computer simulated using nine weights, each adapted by the LMS algorithm. The desired signal is a 10 Hz sinewave sampled at $f = 128$ Hz. The noise input is the sum of two equal sinewaves with frequency $f = 10$ Hz and frequency $f = 35$ Hz. The number of samples is $N = 128$ and the adaptation constant is $\mu = K = 0.1$. The system output

converges after a short learning period to the desired signal. Figures 4.1 through 4.3 depict these results.

A sample of three of the nine weights (w_0, w_4, w_8) is shown in Figures 4.4 through 4.6 to illustrate the steady state solution of the adaptive weights. After just N = 24 all weights have settled to within 0.1% of their final value. Figure 4.7 is the system error. It should be noted that a faster sampling frequency of f = 256 reduces this error even further.

The next simulation is similar to the first with the exception that this system uses only four weights instead of the nine weights used previously. Again the noise input is the sum of two equal sinewaves (f = 4 Hz, f = 20 Hz) and the reference or desired waveform is a sinewave at f = 4 Hz. The adaptation constant is K1 = 0.1, N = 128 and the sampling frequency for the inputs is f = 128 Hz. Figures 4.8 and 4.9 show the filter input and the reference signal. It is evident from the system output shown in Figure 4.10 that this system does not track as well as the previous system with nine weights. Plotting weight #1 of the system through time, as shown in Figure 4.11, illustrates the oscillatory nature of the weight values.

The same unconstrained simulation is repeated with the noise input taken as zero mean uncorrelated noise (generated by the HP-85 random number generator) summed with a sinewave of frequency f = 10 Hz. Figure 4.12 shows the noise input.

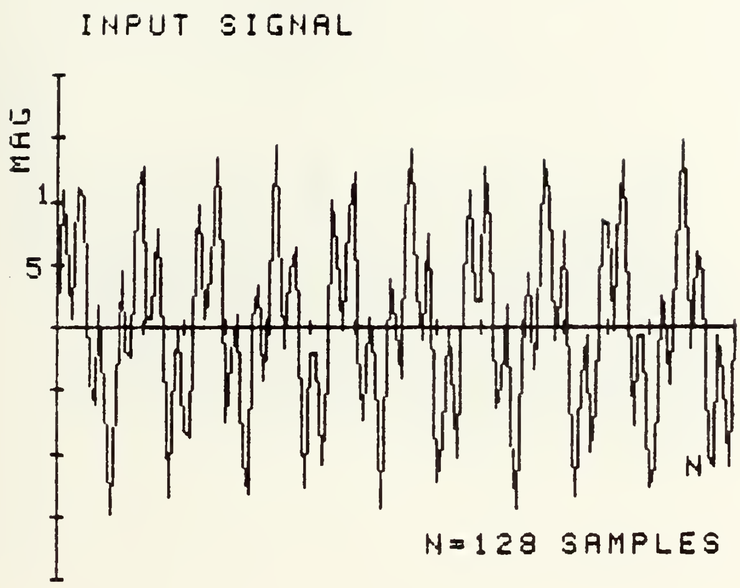

Fig. 4.1. System Input - Nine Adaptive Non-Constrained Weights

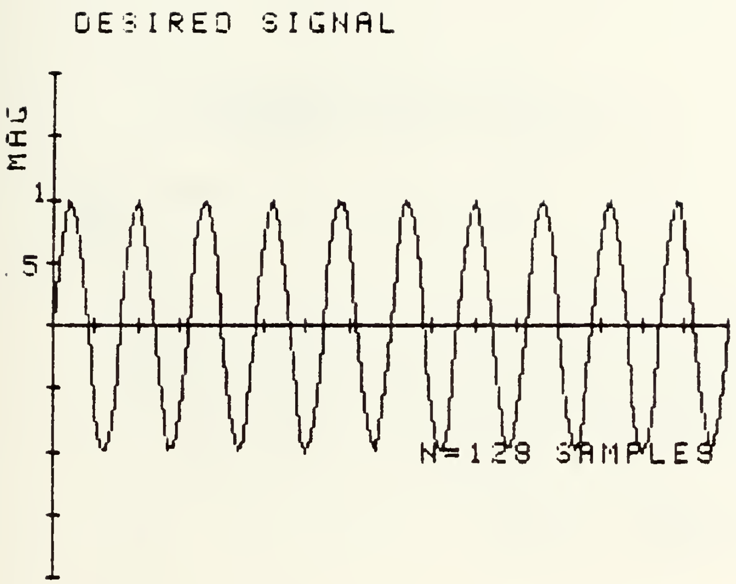

Fig. 4.2. System Reference Signal
Sampling Frequency F_s = 128 Hz

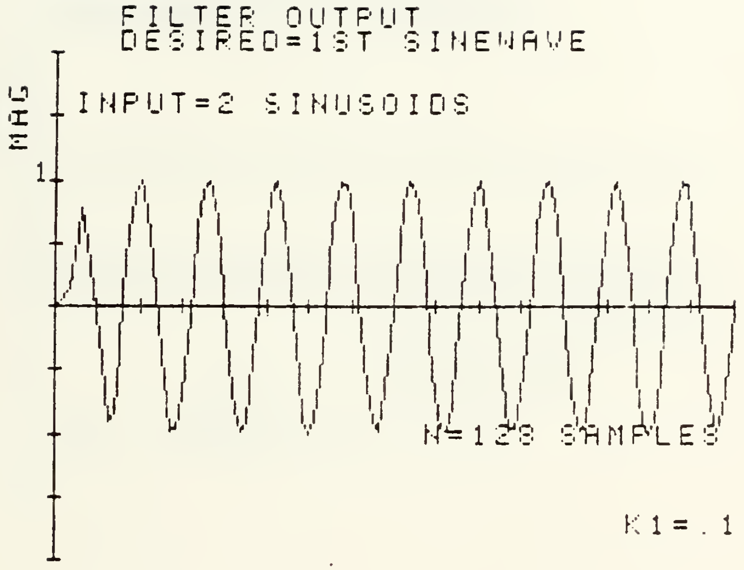

Fig. 4.3. System Output - Nine Adaptive Weights

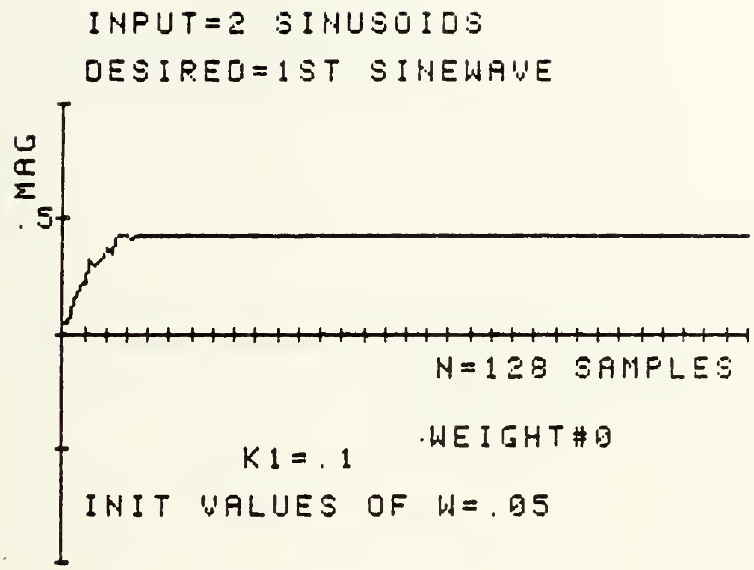

Fig. 4.4. Weight #0 Plotted Through N = 128 Samples
Nine Adaptive Weights

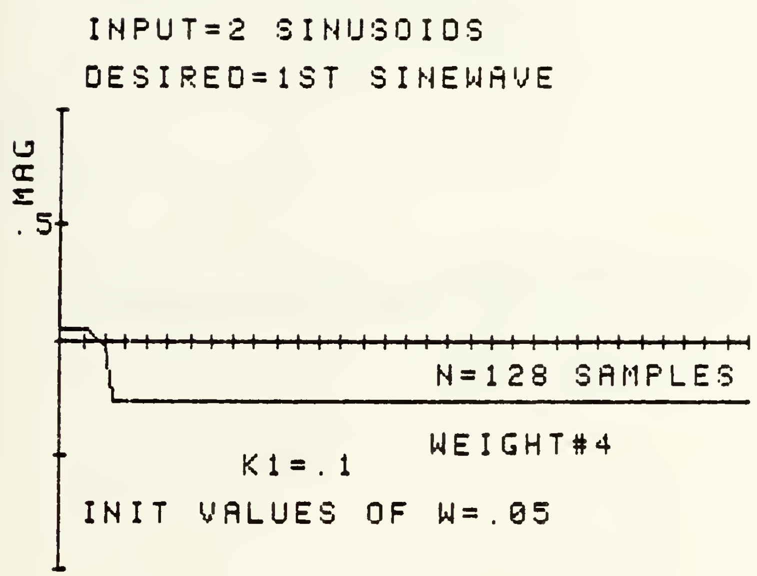

Fig. 4.5. Weight #4 Plotted Through N = 128 Samples
Nine Adaptive Weights

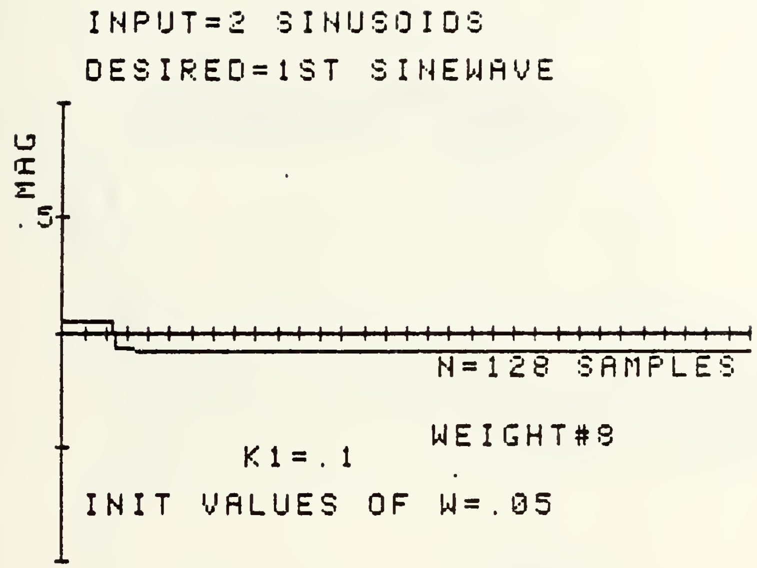

Fig. 4.6. Weight #8 Plotted Through N = 128 Samples
Nine Adaptive Weights

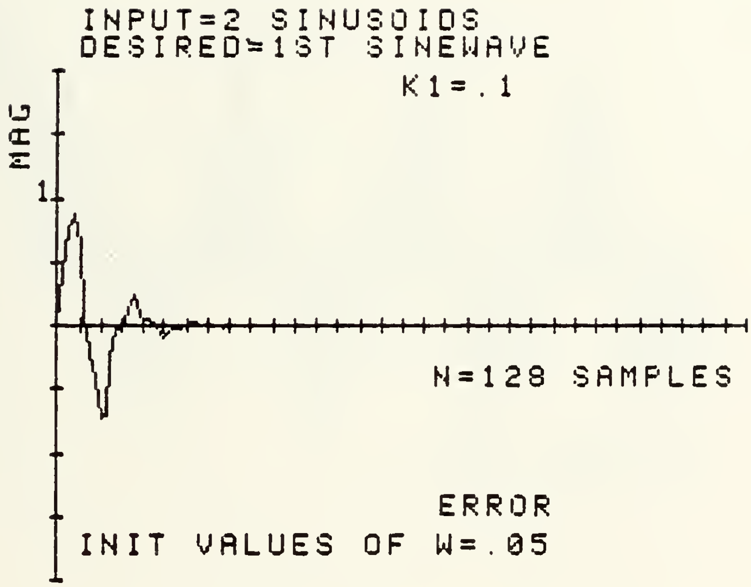

Fig. 4.7. System Error - Nine Adaptive Weights

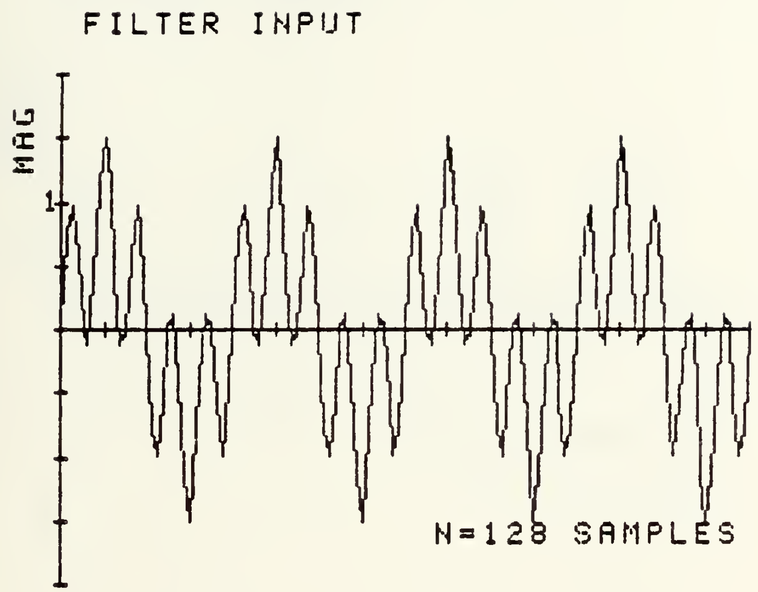

Fig. 4.8. System Noise Input - Four Adaptive Weights

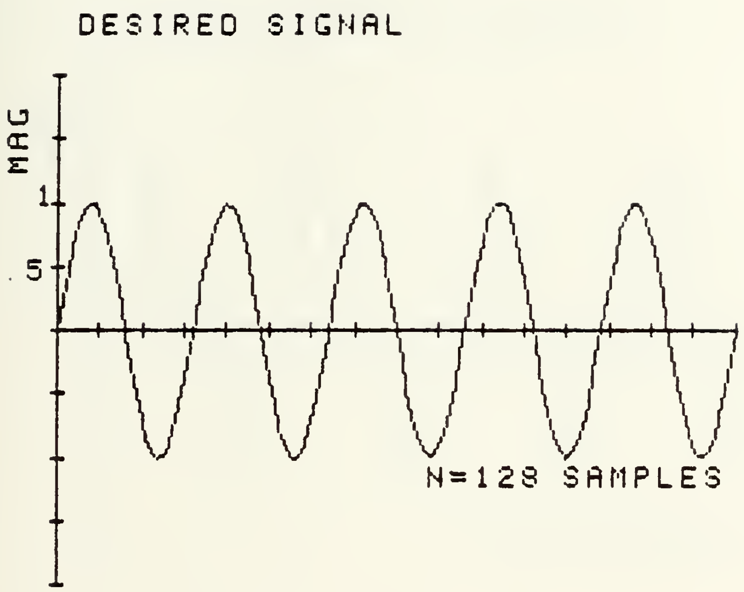

Fig. 4.9. System Reference

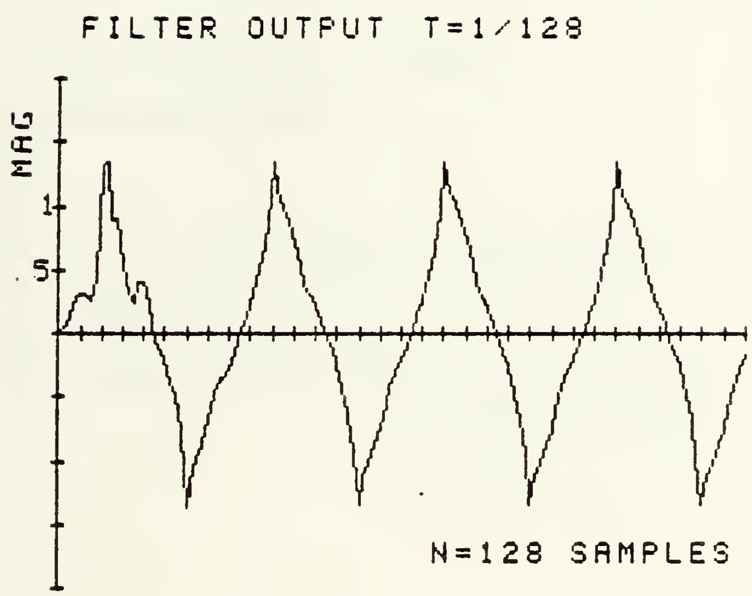

Fig. 4.10. System Output - Four Adaptive Weights

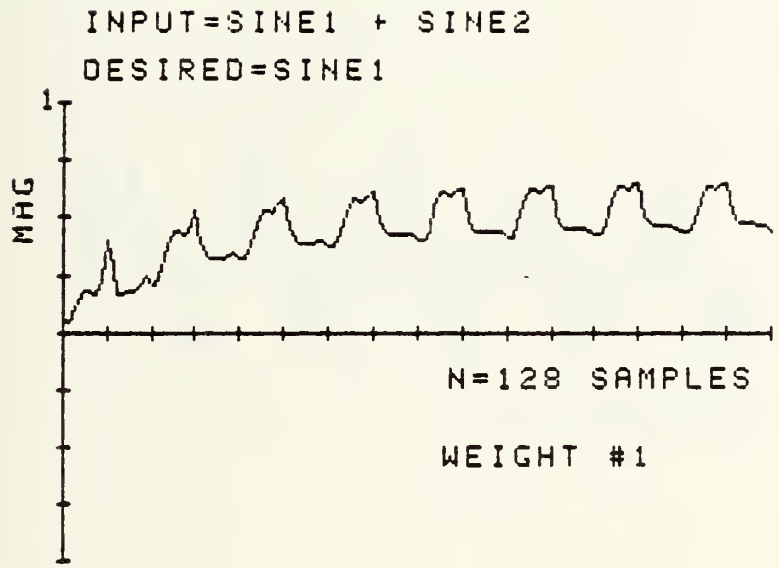

Fig. 4.11. Weight #1 Plotted Through N = 128 Samples

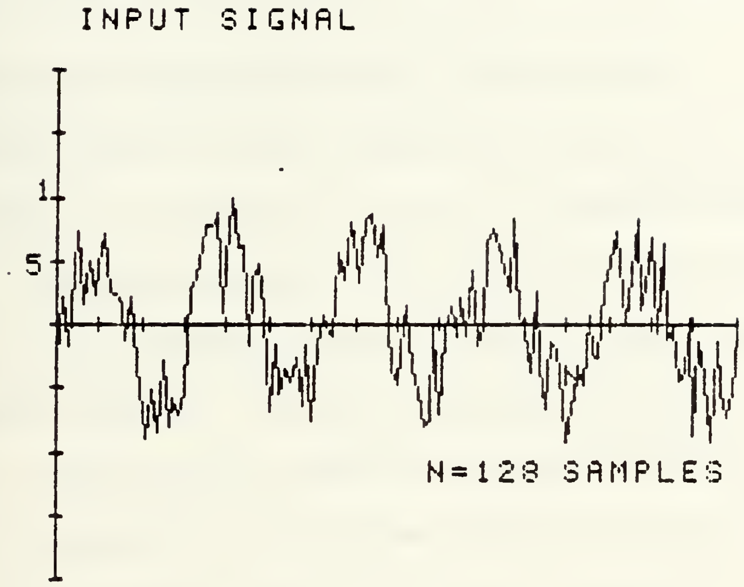

Fig. 4.12. System Input - S/N = 1.64 dB

The desired signal is the sinewave shown in Figure 4.13. Nine adaptive weights are utilized. The system sampling frequency is f = 256 Hz and the adaptation constant is K1 = 0.1. The signal-to-noise ratio of the input signal is 1.64 dB. Figure 4.14 shows the system output while Figure 4.15 illustrates the system error. Figure 4.16 shows one of the system weights plotted through time.

B. CONSTRAINED ADAPTIVE SYSTEM SIMULATIONS

The simulation of the angularly constrained adaptive systems shown in Figures 3.1 and 3.3 produces nearly identical results. Computationally, the iterative solution required for the direct implementation is much slower than that of the cascaded system. Since both system outputs are virtually identical, those of the faster cascaded system are shown. All of the linearly constrained adaptive results utilize nine adaptive weights and all of the angularly constrained systems utilize four weights.

Using the cascaded adaptive system shown in Figure 3.3 simulation results are obtained for a noise input shown in Figure 4.17 of two sinewaves of frequencies f = 4 Hz and f = 20 Hz. The sampling frequency f = 128 Hz. The desired signal shown in Figure 4.18 is a sinewave of frequency f = 4 Hz. The adaptation constant is K1 = 0.1. The zeros are desired at a radius of 1 and with a separation of 12 degrees. The system output tracks the desired signal

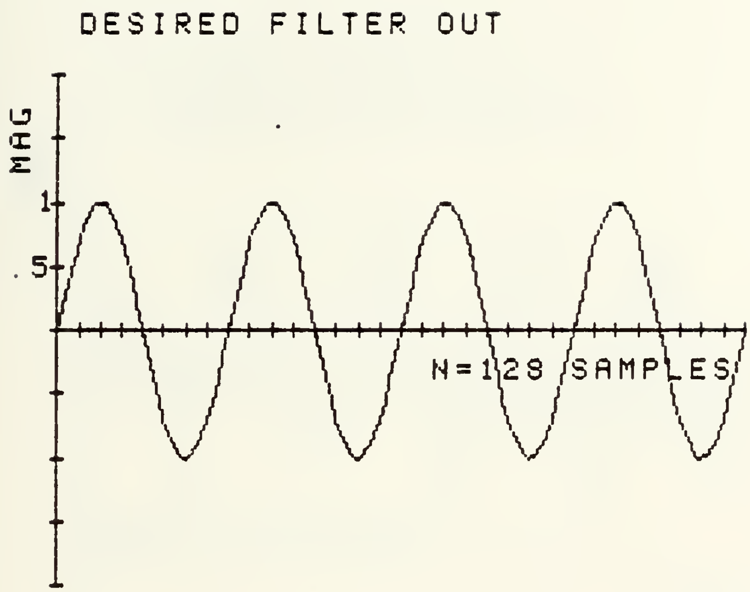

Fig. 4.13. System Reference Signal

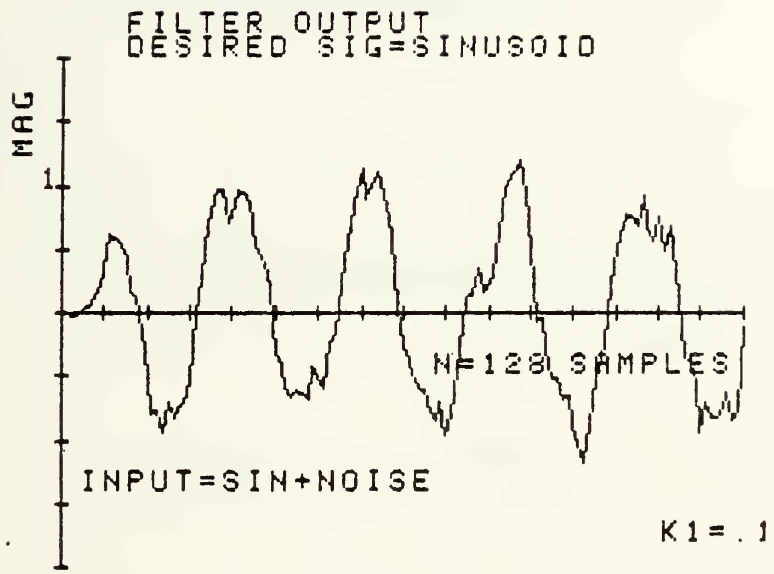

Fig. 4.14. System Output - Four Adaptive Weights

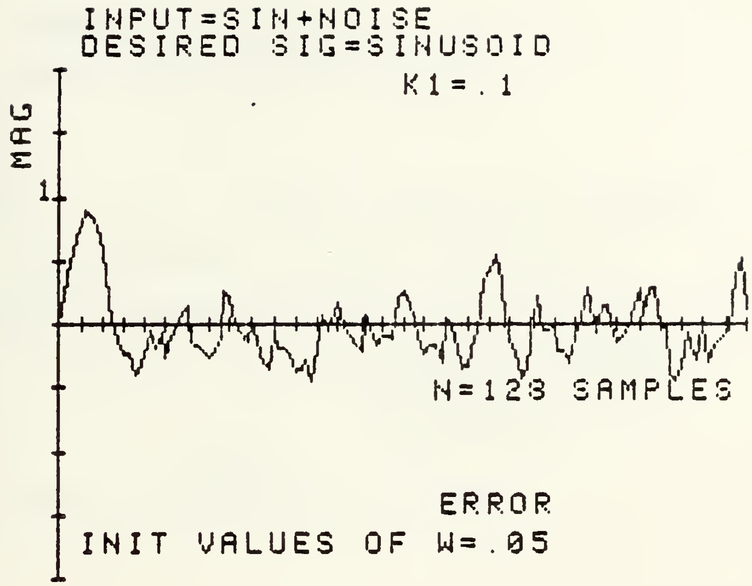

Fig. 4.15. System Error - Four Adaptive Weights

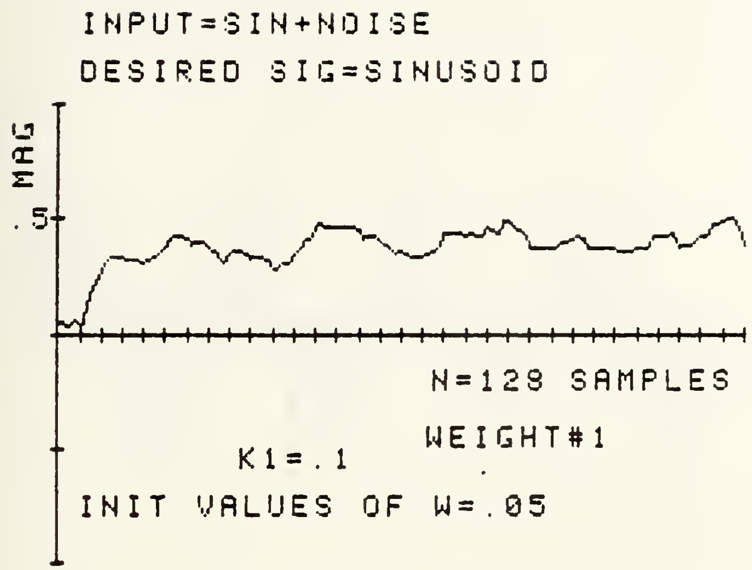

Fig. 4.16. Weight #1 Plotted Through N = 128 Samples
Four Adaptive Weights

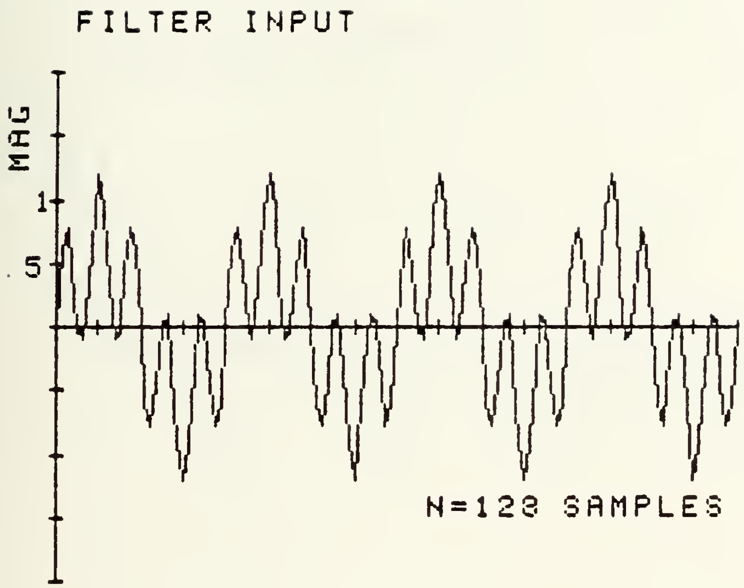

Fig. 4.17. System Input - Sum of Two Sinewaves

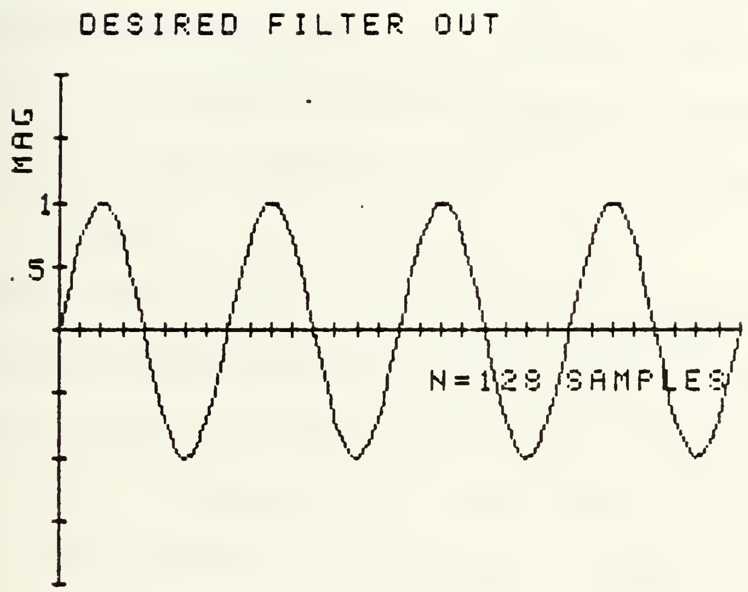

Fig. 4.18. Reference Input

as shown in Figure 4.19 and places the system zeros with the exact requested degree of separation as shown in Figure 4.20. Zero 1 is at radius 1 and an angle of 63.39 degrees. Zero 2 is at a radius of 1 and an angle of 51.39 degrees. The other two zeros are the conjugate pairs of the first two.

The adapted weight number 1 is plotted through time in Figure 4.21. The same oscillating pattern is shown here as is depicted in Figure 4.11. It is felt that, as in the case of Figure 4.11, using a greater number of weights would eliminate the noisy weight pattern and accordingly produce a cleaner output.

Suppose now that the user wishes to drastically alter the frequency characteristics. The input to the system remains the same (as shown in Figure 4.17) however now the desired angle of separation of the zeros is 130 degrees instead of the original 12 degrees. System output is shown in Figure 4.22, the reference signal is shown in Figure 4.23, and the z-plane zero location is shown in Figure 4.24. Zero number 1 is at a radius of 1 and at an angle of 158.55 degrees and zero number 2 is at a radius of 1 and an angle of 28.55 degrees, providing the prescribed 130 degrees of separation. The remaining two zeros form the conjugate pair. The sampling frequency is changed to f = 256 Hz and the adaptation constant K1 is changed to 0.05.

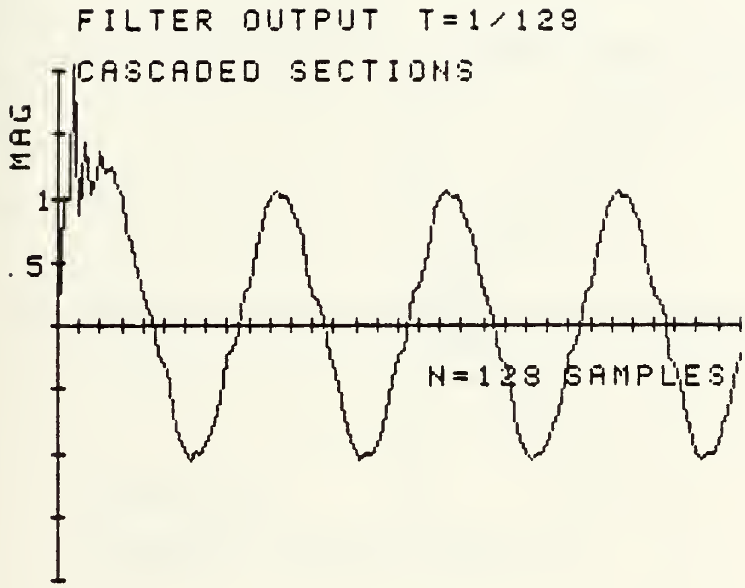

Fig. 4.19. System Output - Four Adaptive Weights
Desired Zero Separation = 12 Degrees

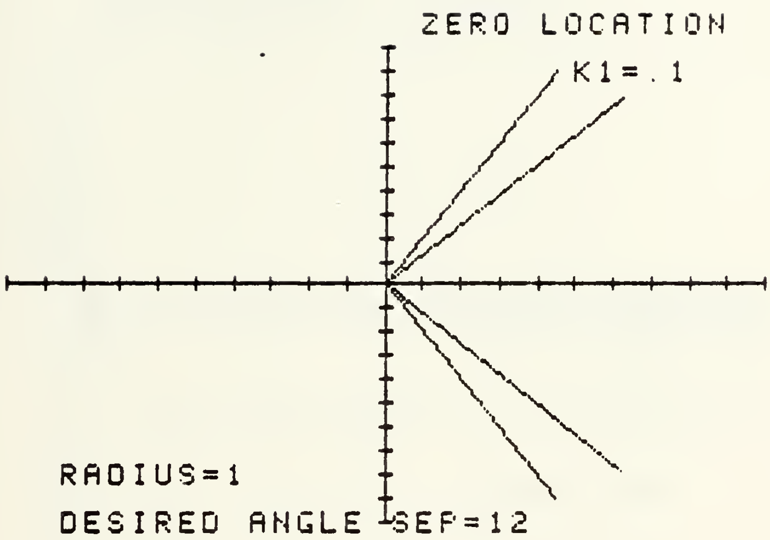

Fig. 4.20. Z-Plane Zero Location - Four Adaptive Weights

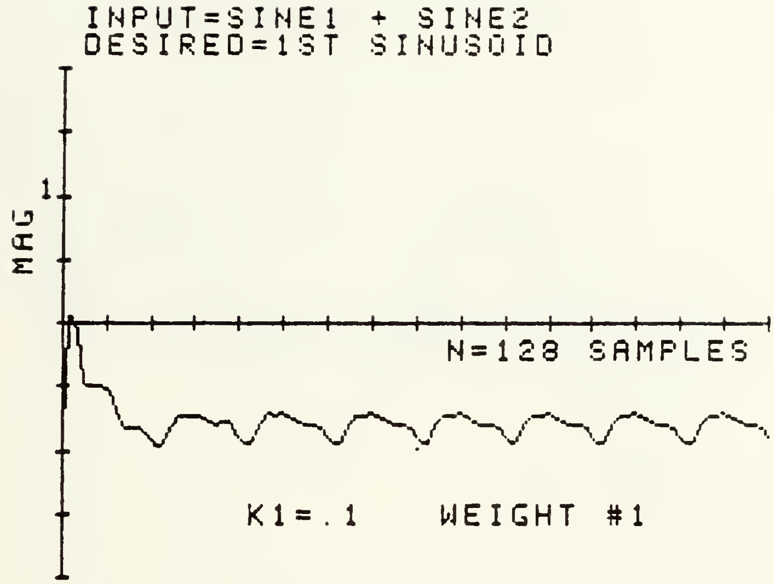

Fig. 4.21. Weight #1 Plotted Through N = 128 Samples
Four Adaptive Weights

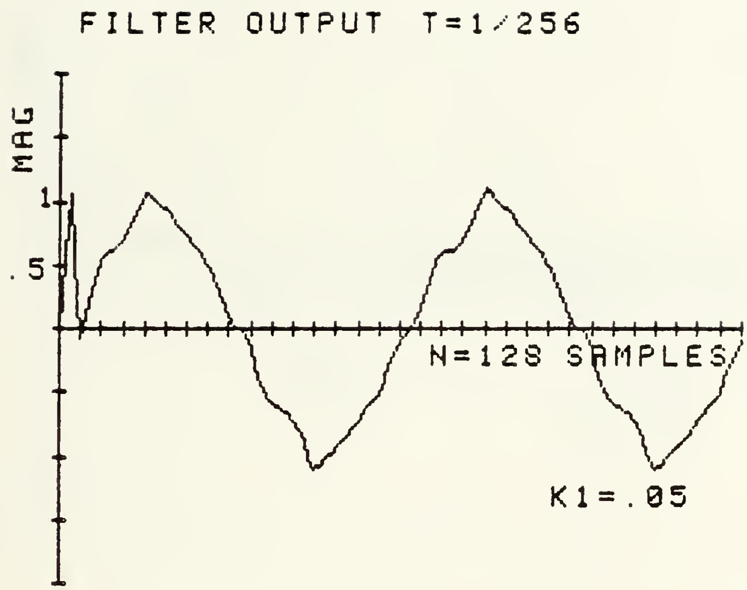

Fig. 4.22. Filter Output - Four Adaptive Weights

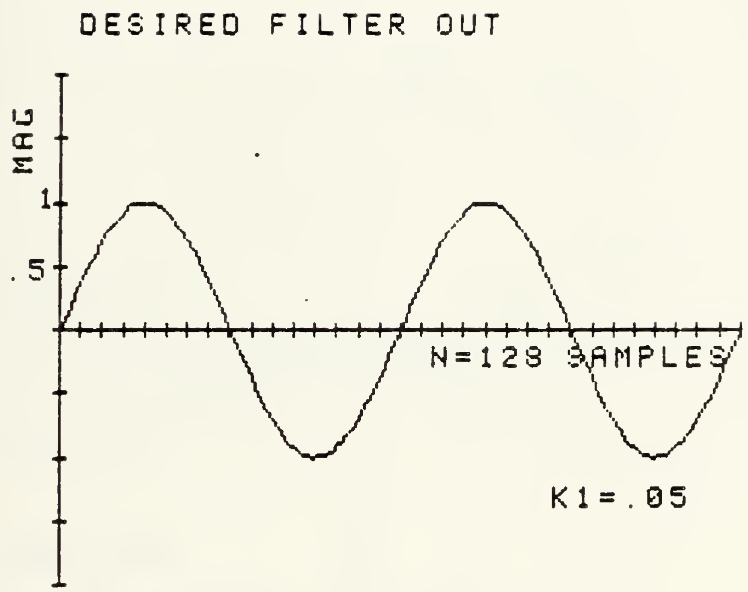

Fig. 4.23. Reference Signal

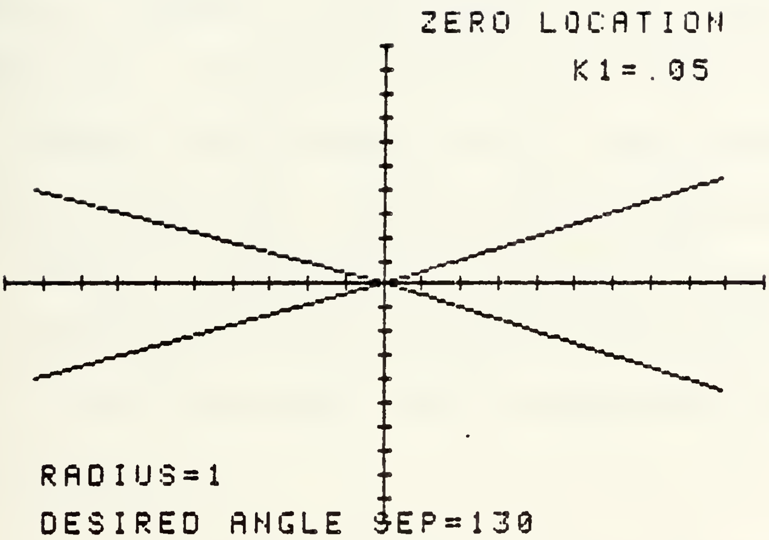

Fig. 4.24. Z-Plane Zero Location

The same simulations are performed for two sinusoids of frequencies f = 10 Hz and f = 35 Hz. A sampling frequency of f = 256 Hz is used with an adaptation constant K1 = 0.1. Figures 4.25 through 4.28 depict the system input (the sum of the two sinusoids), the filter output, the desired signal, and the z-plane zero locations for a desired separation of 20 degrees. Zero 1 is located at a radius of 1 and an angle of 50.2 degrees. Zero 2 is located at a radius of 1 and an angle of 70.2 degrees. The remaining two zeros are the conjugate pairs. Figures 4.29 and 4.30 show the system output and z-plane zero location for the same system input and same reference but with a desired zero separation of 40 degrees. Zero 1 is located at a radius of 1 and an angle of 45.43 degrees and zero 2 is located at a radius of 1 and an angle of 85.43 degrees. The remaining two zeros are conjugate pairs.

C. LINEARLY CONSTRAINED ADAPTIVE FILTER SIMULATIONS USING THE LAGRANGE MULTIPLIER TECHNIQUE

The linearly constrained LMS adaptive filter shown in Figure 3.4 is simulated for various values of K, the linear constraint value. Two types of system inputs are used--random "white" noise and a sinusoidal signal of frequency 4 Hz. In each case the desired signal is a damped exponential with a final value of 0.2.

Figure 4.31 depicts the filter output superimposed upon the desired output of a damped exponential of final value

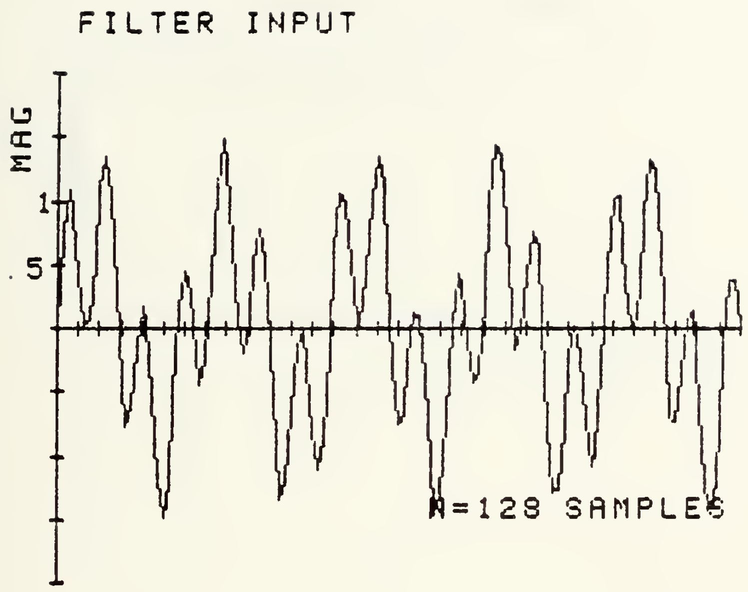

Fig. 4.25. System Input - Four Adaptive Weights

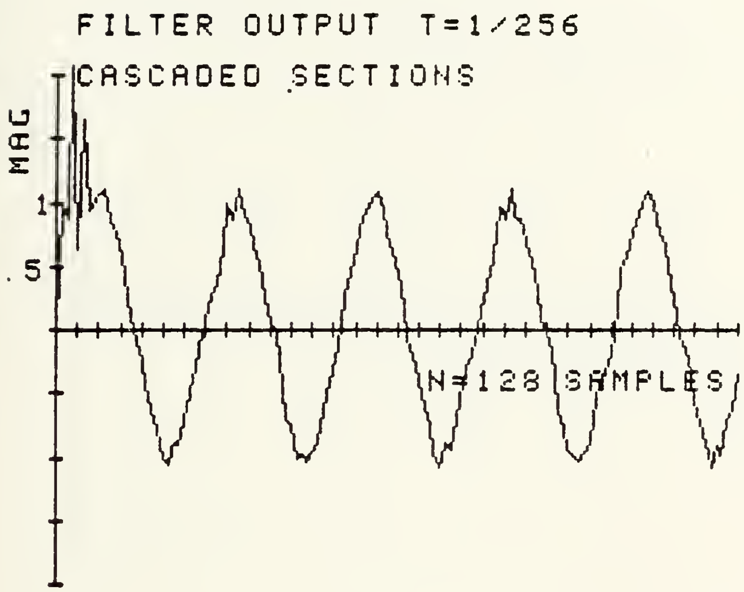

Fig. 4.26. System Output - Four Adaptive Weights
Desired Zero Separation = 20 Degrees

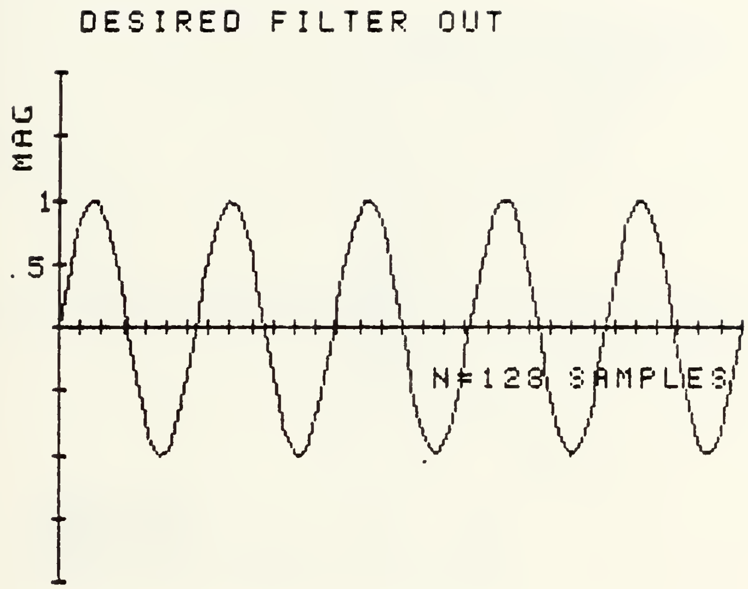

Fig. 4.27. Reference Signal

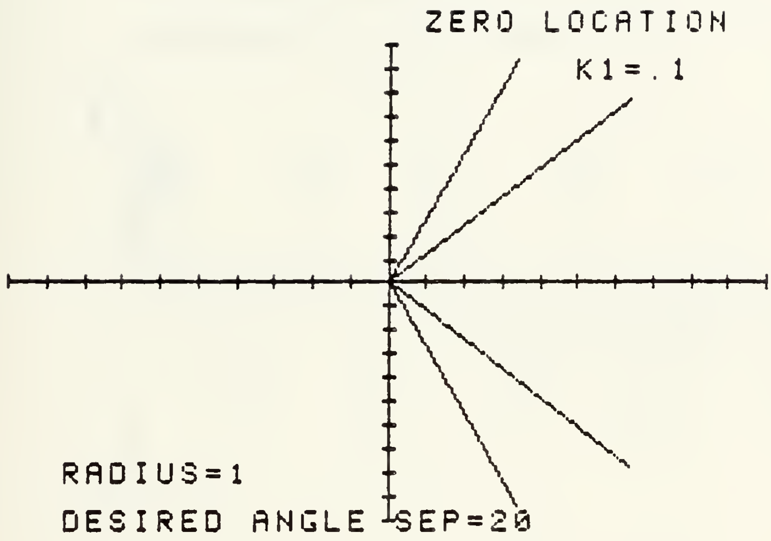

Fig. 4.28. Z-Plane Zero Locations - Four Adaptive Weights

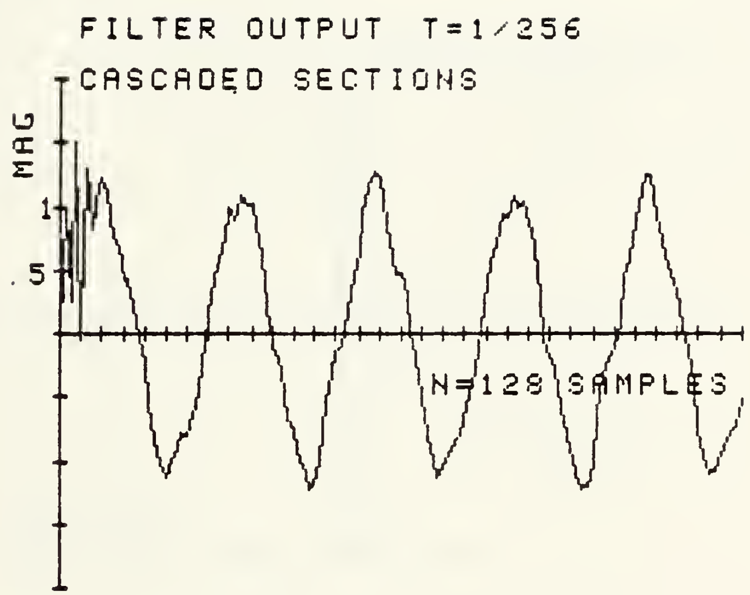

Fig. 4.29. System Output - Four Adaptive Weights
Desired Zero Separation = 40 Degrees

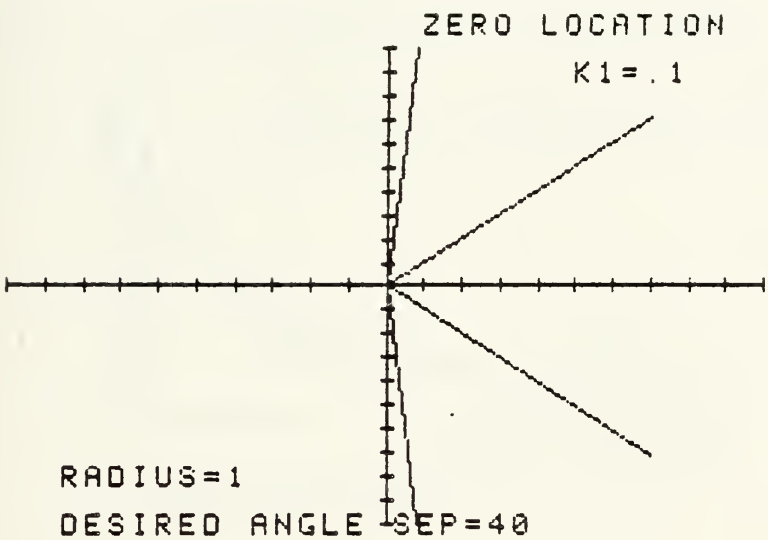

Fig. 4.30. Z-Plane Zero Location

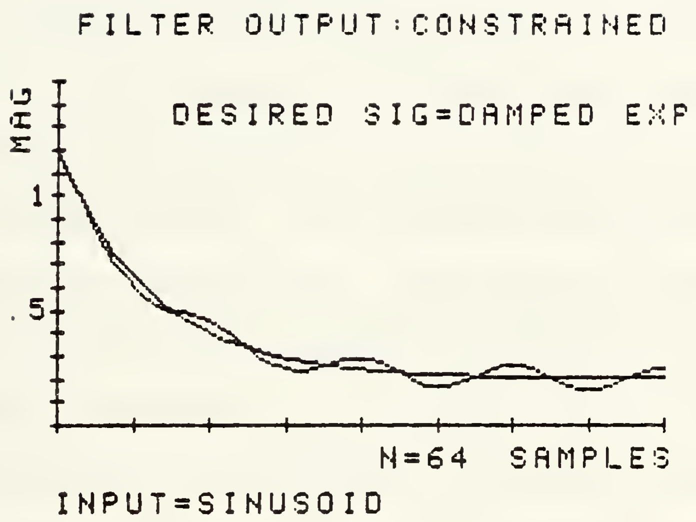

Fig. 4.31. System Output and Desired Signal
Linear Constraint K = 0

0.2. The linear constraint K is arbitrarily chosen to be zero. The adaptation constant is K1 = 1 E-6. The system error is shown in Figure 4.32. The sum of the weights plotted through time is shown in Figure 4.33. The final value of the Lagrange multiplier, that is $\lambda(N)$ at time N = 64, is found to be -1.022 E-4. The final value of the sum of weights is found to be 0.059 compared with a requested value of zero.

The same system is again simulated for a linear constraint of K = -0.2. Figure 4.34 shows that the sinusoidal input converges to the desired damped exponential of final value 0.2. Interestingly, the response converges and then begins to build an oscillation. This type of response in adaptive systems has been observed elsewhere and may be due to an arithmetic precision problem [Ref. 10]. Figure 4.35 shows the system error and Figure 4.36 shows the sum of weights graph plotted through time. The final value of the sum of weights is found to be -0.2427 compared with the desired value of -0.2.

The simulation is again performed, now with a "white" noise input and the same damped exponential with final value of 0.2 as the desired signal. In the first of the three noise input simulations the sum of weights constraint K is set to -0.1; the second simulation has K = 0.2; the final simulation uses K = 0.5. Figures 4.37 through 4.42 depict the simulation results. In Figure 4.42 the sum of the

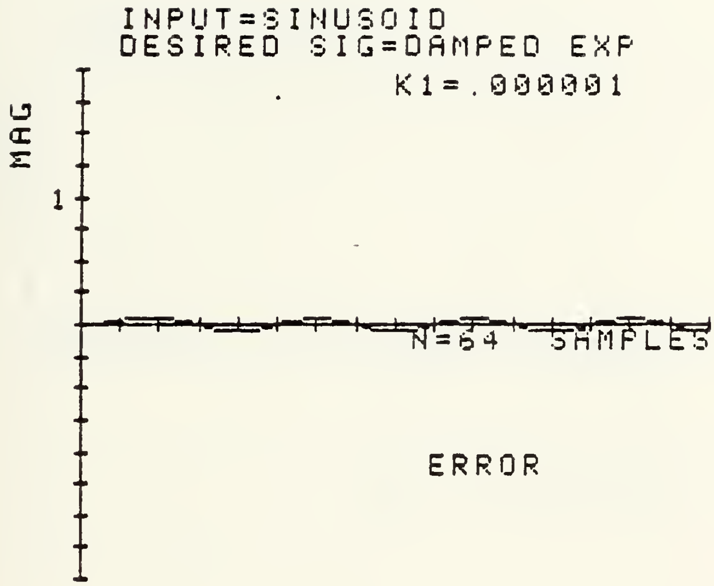

Fig. 4.32. System Error y(n)

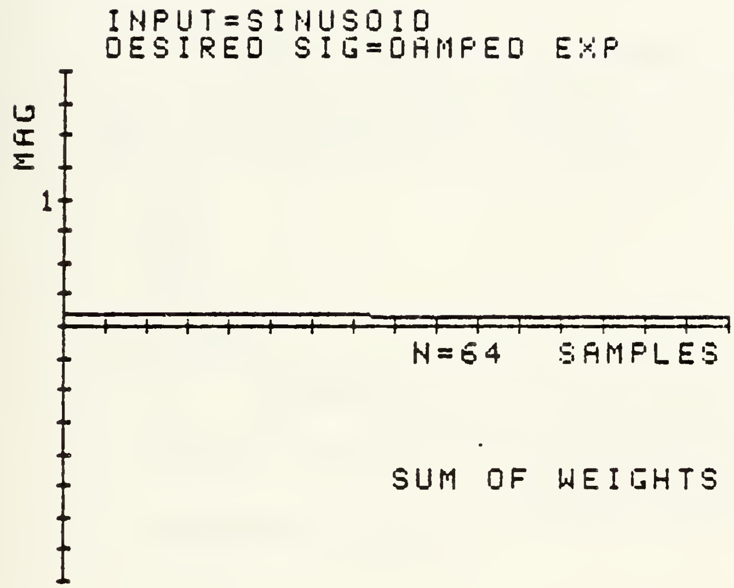

Fig. 4.33. Sum of Weights Plotted Through N = 128 Samples. Constraint K = 0

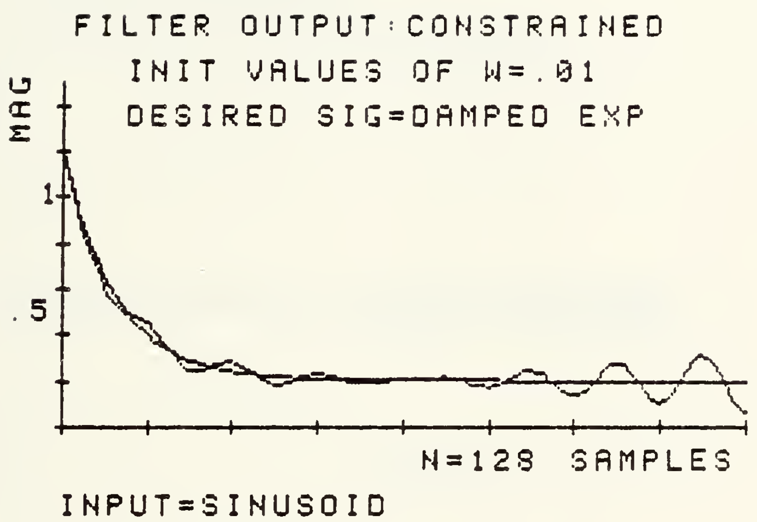

Fig. 4.34. System Output and Desired Signal
K = -.2

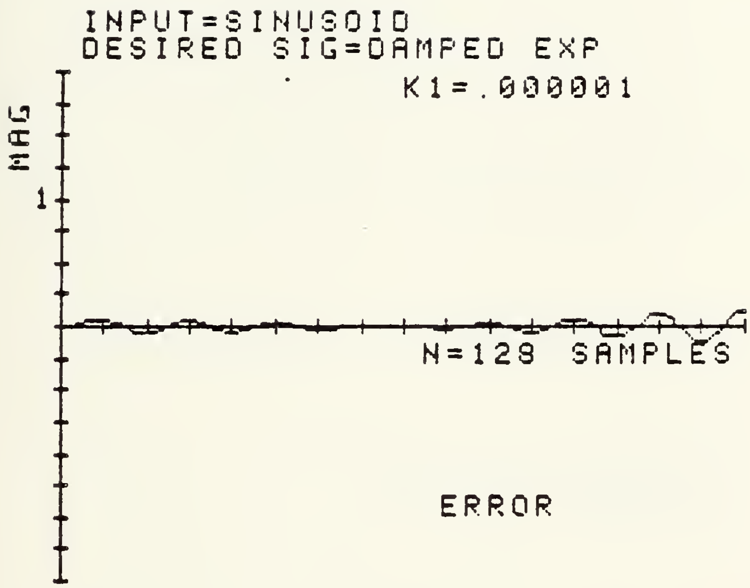

Fig. 4.35. System Error
Linear Constraint K = -.2

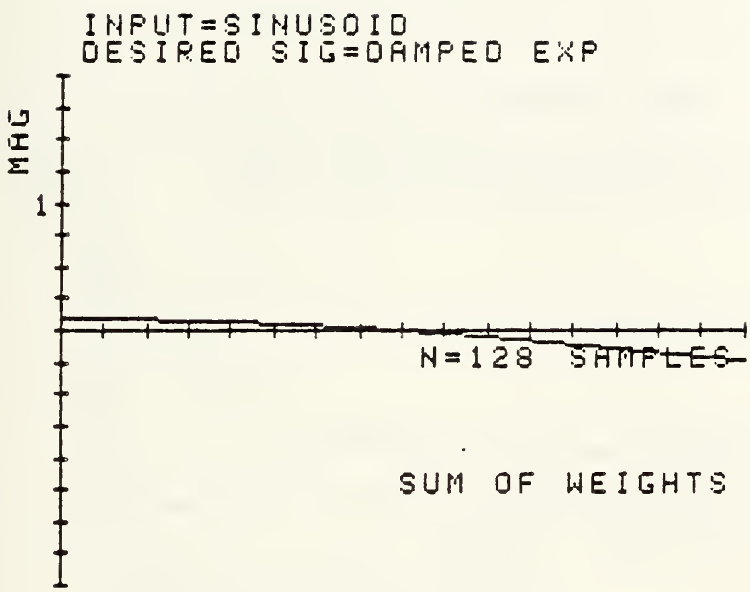

Fig. 4.36. Sum of Weights Plotted Through N = 128 Samples. Linear Constraint K = -.2

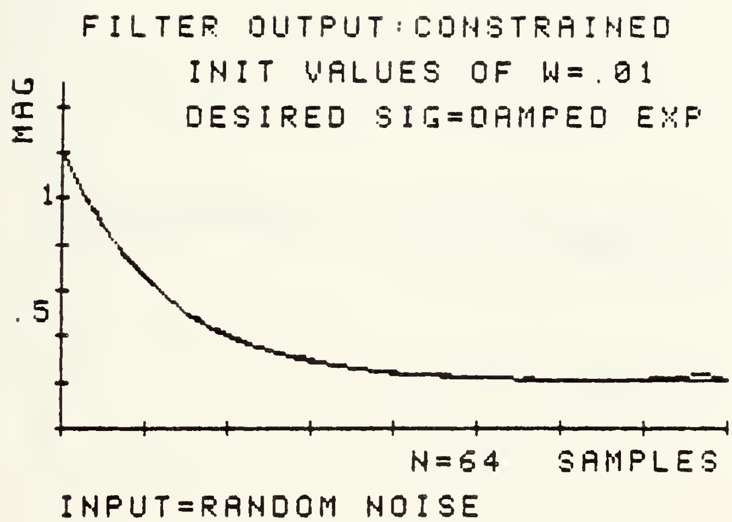

Fig. 4.37. System Output Superimposed over Desired Signal of Damped Exponential K = -.1. Nine Adaptive Weights

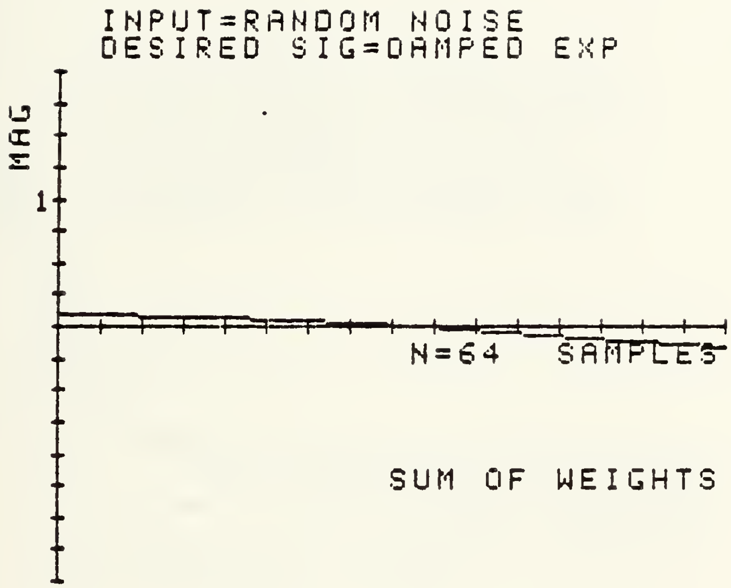

Fig. 4.38. Sum of Weights Plotted Through N = 64 Samples. Constraint K = -0.1 Nine Adaptive Weights

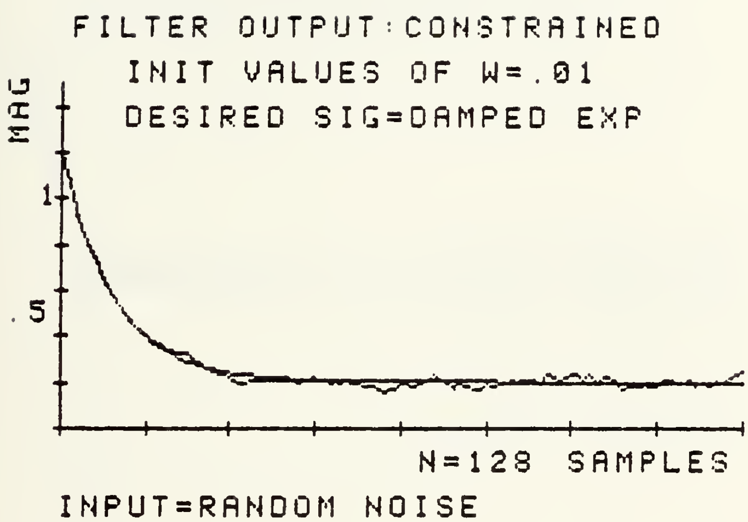

Fig. 4.39. System Output Superimposed over Desired Signal of Damped Exponential K = +0.2
Nine Adaptive Weights
K1 = .00001 K2 = .00008

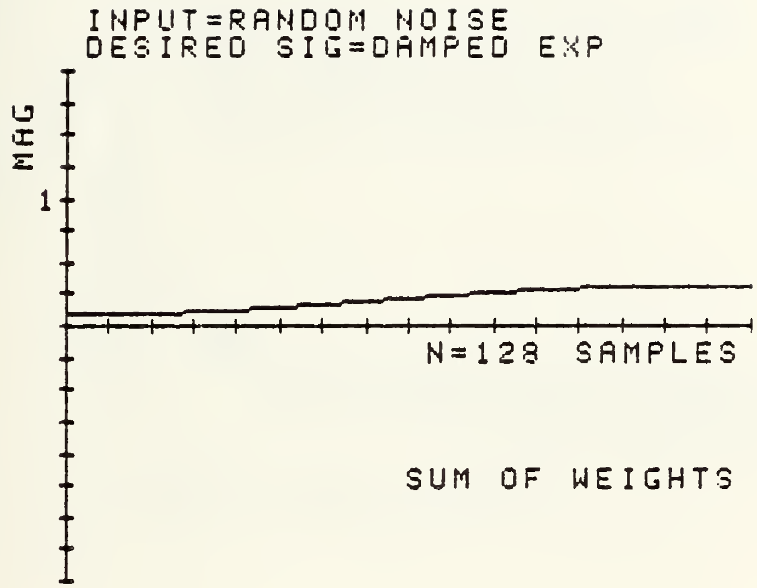

Fig. 4.40. Sum of Weights Plotted for N = 128 Samples
Nine Adaptive Weights. K = .2
K1 = .00001 K2 = .00008

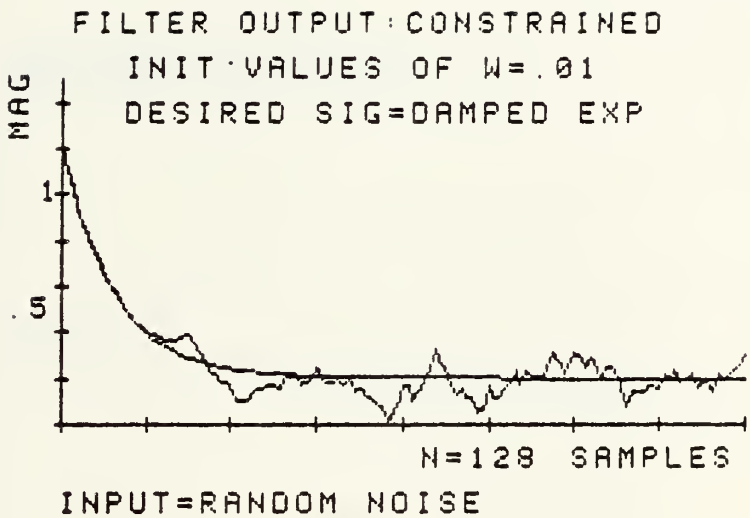

Fig. 4.41. System Output Superimposed over the Desired Signal of a Damped Exponential
K = .5
K1 = .00001 K2 = .00009

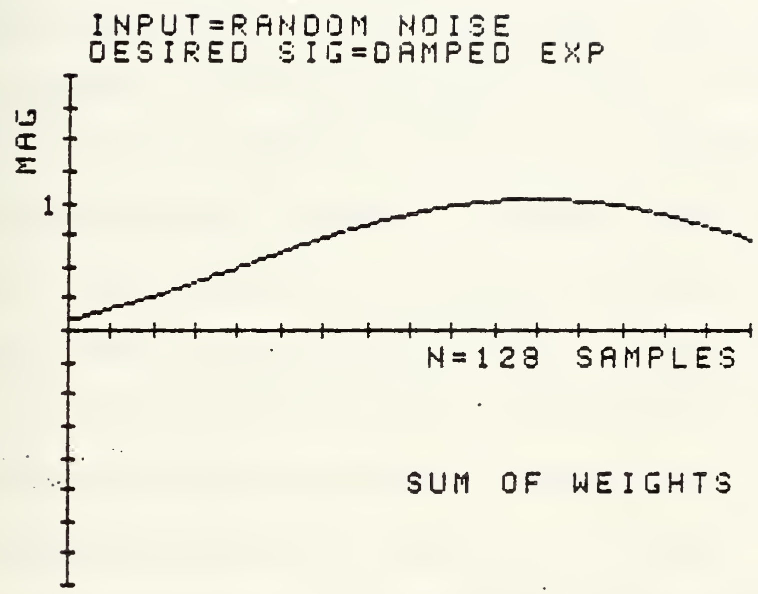

Fig. 4.42. Sum of Weights Plotted Through N = 128 Samples. K = .5
K1 = .00001 K2 = .00009

weights overshoots the constraint of K = 0.5 but returns to within 10% of that value at N = 128.

D. CONCLUSIONS

This thesis has investigated adaptive noise cancellation techniques in detail with the objective of implementing constrained adaptive filters. Two basic types of constrained adaptive filters are presented--the fixed angular separation (direct and cascaded implementations) and the adaptive Lagrangian multiplier linear constraint approach. All three methods produce the desired output within a very close tolerance for the examples tested. As expected, the direct implementation takes longer to converge than the cascade approach. The results are promising, however, several key questions remain to be investigated. In particular the convergence properties of these constrained adaptive algorithms need to be studied in detail. In the case of the Lagrangian multiplier approach, if there is an optimal steady-state solution, the modified LMS algorithm should find a minimum since it is a simplified gradient technique. However the LMS minimum may not be global. For the case of the angular constraint process, the stability of the adaptive algorithm and its convergence also need to be examined, although the master-slave would seem to contain inherent stability.

Finally, the concepts of constrained adaptive algorithms need to be examined with more sophisticated examples using a larger number of weights.

LIST OF REFERENCES

1. Wiener, N., Extrapolation, Interpolation and Smoothing of Stationary Time Series, with Engineering Applications, Wiley, 1949.

2. Kalman, R. and Bucy, R., "New Results in Linear Filtering and Prediction Theory," Trans. ASME, serial D, vol. 83, pp. 95-107, December 1961.

3. Widrow, B., Adaptive Filters, A chapter from Aspects of Network and Systems Theory. Edited by R. E. Kalman and N. DeClaris, Holt, Rinehard and Winston, Inc., N.Y., 1970.

4. Widrow, B. and others, "Adaptive Noise Cancelling: Principles and Applications," Proc. IEEE, vol. 63, pp. 1692-1716, December 1975.

5. McCool, J.M. and Widrow, B., Principles and Applications of Adaptive Filters: A Tutorial Review, Naval Undersea Center, San Diego, Ca., and Stanford University, Stanford, Ca., MUC TP 530 AD-A037755, March 1977.

6. McCool, J.M., The Basic Principles of Adaptive Systems with Various Applications, Naval Undersea Center, San Diego, Ca., Fleet Engineering Department Electronic Division, September 1972.

7. Sandia Labs, Albuquerque, New Mexico, SAND 76-0360, The Adaptive Digital Filter, Glenn R. Elliot and others, Aug. 1976.

8. Frost, O.L. III, "An Algorithm for Linearly Constrained Adaptive Array Processing," Proc. IEEE, Vol. 60, Aug. 1972.

9. Thomas, G.B. Jr., and Finney, R.L., Calculus and Analytic Geometry, 5th ed., pp. 620-627, Addison-Wesley, 1979.

10. Gibson, C.J. and Haykin, S., "Learning Characteristics of Adaptive Lattice Filtering Algorithms," Trans. on ASSB, Vol. 28, December 1980.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Technical Information Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0142 Naval Postgraduate School Monterey, California 93940	2
3.	Department Chairman, Code 62 Department of Electrical Engineering Naval Postgraduate School Monterey, California 93940	1
4.	Prof. S. Parker, Code 62Px Department of Electrical Engineering Naval Postgraduate School Monterey, California 93940	3
5.	Prof. S. Jauregui, Code 62Ja Department of Electrical Engineering Naval Postgraduate School Monterey, California 93940	1
6.	Lieutenant M. H. Davis, USN 496 Mar Vista Drive Monterey, California 93940	1

Thesis
D17232 Davis 193222
c.1
Constrained adaptive
least mean squared
filters.